Journal of Translational Medicine

RESEARCH

Open Access

A comparative study on ctDNA and tumor DNA mutations in lung cancer and benign cases with a high number of CTCs and CTECs

Jianzhu Xie[1†], Binjie Hu[1†], Yanping Gong[1†], Sijia He[1], Jun Lin[2], Qian Huang[1*] and Jin Cheng[1*]

Abstract

Background Liquid biopsy provides a non-invasive approach that enables detecting circulating tumor DNA (ctDNA) and circulating tumor cells (CTCs) using blood specimens and theoretically benefits early finding primary tumor or monitoring treatment response as well as tumor recurrence. Despite many studies on these novel biomarkers, their clinical relevance remains controversial. This study aims to investigate the correlation between ctDNA, CTCs, and circulating tumor-derived endothelial cells (CTECs) while also evaluating whether mutation profiling in ctDNA is consistent with that in tumor tissue from lung cancer patients. These findings will help the evaluation and utilization of these approaches in clinical practice.

Methods 104 participants (49 with lung cancer and 31 with benign lesions) underwent CTCs and CTECs detection using integrating subtraction enrichment and immunostaining-fluorescence in situ hybridization (SE-iFISH) strategy. The circulating cell-free DNA (cfDNA) concentration was measured and the mutational profiles of ctDNA were examined by Roche AVENIO ctDNA Expanded Kit (targeted total of 77 genes) by next generation sequencing (NGS) in 28 patients (20 with lung cancer and 8 with benign lesions) with highest numbers of CTCs and CTECs. Mutation validation in matched tumor tissue DNA was then performed in 9 patients with ctDNA mutations using a customized xGen pan-solid tumor kit (targeted total of 474 genes) by NGS.

Results The sensitivity and specificity of total number of CTCs and CTECs for the diagnosis of NSCLC were 67.3% and 77.6% [AUC (95%CI): 0.815 (0.722–0.907)], 83.9% and 77.4% [AUC (95%CI): 0.739 (0.618–0.860)]. The concentration of cfDNA in plasma was statistically correlated with the size of the primary tumor ($r=0.430$, $P=0.022$) and CYFRA 21–1 ($r=0.411$, $P=0.041$), but not with the numbers of CTCs and CTECs. In this study, mutations were found to be poorly consistent between ctDNA and tumor DNA (tDNA) in patients, even when numerous CTCs and CTECs were present.

Conclusion Detection of CTCs and CTECs could be the potential adjunct tool for the early finding of lung cancer. The cfDNA levels are associated with the tumor burden, rather than the CTCs or CTECs counts. Moreover, the poorly consistent mutations between ctDNA and tDNA require further exploration.

Keywords Lung cancer, ctDNA, Mutations, NGS, CTCs, CTECs

†Jianzhu Xie, Binjie Hu and Yanping Gong contributed equally to this work.

*Correspondence:
Qian Huang
huangqian_sjtu@163.com
Jin Cheng
jin.cheng@shgh.cn
Full list of author information is available at the end of the article

Background

Lung cancer ranks second as one of the most common malignant tumors worldwide, with an incidence rate of 31.5 per 100,000 in males and 14.6 per 100,000 in females [1]. In China, the mortality rate is significantly higher than that in the United States (about 1.4 times), even though the incidences are similar [2]. Furthermore, the five-year survival rate is relatively lower, compared with developed countries, which may be caused by the lack of national-scale lung cancer screening and insufficient population coverage [2]. In this regard, effective screening and early diagnosis have increasingly become the keys to reducing the mortality of lung cancer.

Tissue biopsy remains the gold standard but can be restricted by implementation conditions such as unavailable or inadequate tumor tissues and poor compliance [3]. Currently, low-dose computed tomography (LDCT) scanner is the primary screening method for early-stage lung cancer. Numerous large-scale clinical trials have demonstrated a significant reduction in mortality rates accompanying routine LDCT screening [4, 5]. However, LDCT is limited by the high false positive rate and unnecessary radiation exposure and there is still a need for further confirmation by pathological diagnosis. Consequently, LDCT is not utilized as a standalone method for lung caner screening in the majority of European countries [6]. Therefore, non-invasive and effective tools for lung cancer diagnosis are urgently needed.

Multiple biomarkers in the peripheral blood such as shed tumor cells and cancer-derived molecules have been explored by scientists in order to provide more valuable information, known as liquid biopsy [6–9]. Among these, circulating tumor cells (CTCs) have sparked considerable interest as a potential biomarker for monitoring cancer presence and progression. Furthermore, a subset of these cells has undergone epithelial-mesenchymal transition (EMT) or other phenotype alterations to evade immune surveillance, potentially contributing to distant metastasis [10, 11]. Detection of CTCs is still challenging due to their low numbers in the bloodstream, which makes it difficult to distinguish them from other cells. Despite these obstacles, researchers have made significant progress in developing new technologies and methods for the isolation and identification of CTCs. One promising development is the creation of microfluidic devices designed to quickly isolate CTCs from blood samples [12, 13]. Other progress involve capturing and recognizing CTCs as well as circulating tumor-derived endothelial cells (CTECs) using tumor cell-specific antibodies and probes, and identifying tumor-related mutated gene or various genetic modifications such as methylation through sequencing or PCR [14–19].

Circulating tumor DNA (ctDNA), which is defined as extracellular DNA molecules either passive released into body fluids by apoptotic, necrotic, pyroptotic tumor cells or active released by tumor cells [20], can be a promising tumor biomarker. The molecular diagnostic test of key tumor gene mutations as well as drug-targeted mutations in patients who failed tissue test has been widely adopted [21–23]. One of the earliest applications of ctDNA detection is to identify epidermal growth factor receptor (*EGFR*) mutations for the guidance of *EGFR* inhibitor therapy in non-small cell lung cancer (NSCLC), which has been approved by the Food and Drug Administration (FDA) [24] and further included in the new College of American Pathologists (CAP)/International Association for the Study of Lung Cancer (IASLC)/Association for Molecular Pathology (AMP) testing guidelines [25]. Additionally, there is evidence to support that methylation profiles of ctDNA can enhance both diagnostic and monitoring capabilities for lung cancer [26–28]. Other researchers also confirmed that the presence of mutated ctDNA after radiotherapy or chemotherapy can predict the existence of minimal residual lesions, which may be directly related to tumor metastasis and recurrence [29, 30]. Nevertheless, studies on the concordance of variant profiles between peripheral blood ctDNA and tumor DNA (tDNA) remain controversial [21, 31]. Furthermore, several hurdles still need to be overcome in ctDNA testing, including low concentration in asymptomatic patients, biological contamination from white blood cells, and variations in sensitivity across different platforms [32].

Although new non-invasion biomarkers for lung cancer have become a research hotspot, the relationship between these novel biomarkers is still not well understood. In this research, we aim to investigate the feasibility of CTCs and CTECs for the diagnosis of NSCLC. We will also examine the coincidence of mutational profiles between ctDNA and tDNA in both lung cancer and benign lung lesion patients, with a particular focus on those with a high number of CTCs and CTECs.

Methods

Subjects and specimens

A total of 104 participants were enrolled in this retrospective study at the Cancer Center in Shanghai General Hospital from August 2017 to April 2019, including 49 cases of newly diagnosed or relapsed non-small cell lung cancer, 31 cases of benign pulmonary disease, and 24 healthy controls. The diagnoses of all the patients were confirmed by clinicians and pathologists. Detailed clinical information is described in Table 1. At first, all the study subjects underwent CTCs and CTECs detection. We further selected the top 20 and 8 patients with the

Table 1 Clinical information of the 104 participating subjects

Characteristics	No.(%)
All Subjects (N = 104)	
Age years, median (range)	58 (37.25–66.75)
Gender (male/female)	51/53
NSCLC (*N* = 49)	
Age years, median (range)	64 (53–69.5)
Gender (male/female)	26/23
Pathological type	
AC	38 (73.4%)
SCC	8 (16.30%)
ASC	3 (6.1%)
Other	2 (4.1%)
TNM stage	
I	27 (55.1%)
II	2 (4.1%)
III	9 (18.4%)
IV	11 (22.5%)
Benign (*N* = 31)	
Age years, median (range)	60 (53–66)
Gender (male/female)	17/14
Healthy (*N* = 24)	
Age years, median (range)	29 (27.25–31.75)
Gender (male/female)	8/16

NSCLC non-small cell lung cancer; *AC* Adenocarcinoma; *SCC* Squamous carcinoma; ASC adenosquamous carcinoma

highest number of CTCs and CTECs from NSCLC and benign cases respectively, and ctDNA mutational profiling detection was carried out in these 28 patients. Of these, ctDNA mutations were detected in 17 patients. Among the 17 patients, 9 patients with paired tissue samples underwent tissue DNA sequencing. Specific exclusion criteria are detailed in Fig. 1. In addition, 4 benign patients with ctDNA variants underwent white blood cell targeted sequencing using 32 Gene and 61 Gene panel. The use of all the specimens from patients was in accordance with the Declaration of Helsinki and approved by the Ethical Review Board of Shanghai General Hospital, Shanghai Jiaotong University School of Medicine, China (2016KY130).

Blood samples were collected using 6 ml ACD anticoagulant vacuum vessel collection tubes for CTCs & CTECs detection and 10 ml EDTA vacuum tubes for ctDNA detection, respectively. All the blood samples were collected before either radiotherapy or chemotherapy. The collected blood samples were mixed gently to avoid any hemolysis.

Tumor DNA was extracted from the FFPE sections using the QIAamp DNA FFPE Tissue Kit (QIAGEN, CA, USA). The thickness of freshly prepared FFPE slices ranged from 5 to 10 μm. Postoperative FFPE sections with tumor cell content > 50% required 2–5 slices. The biopsy FFPE sections with tumor cell content > 20% required

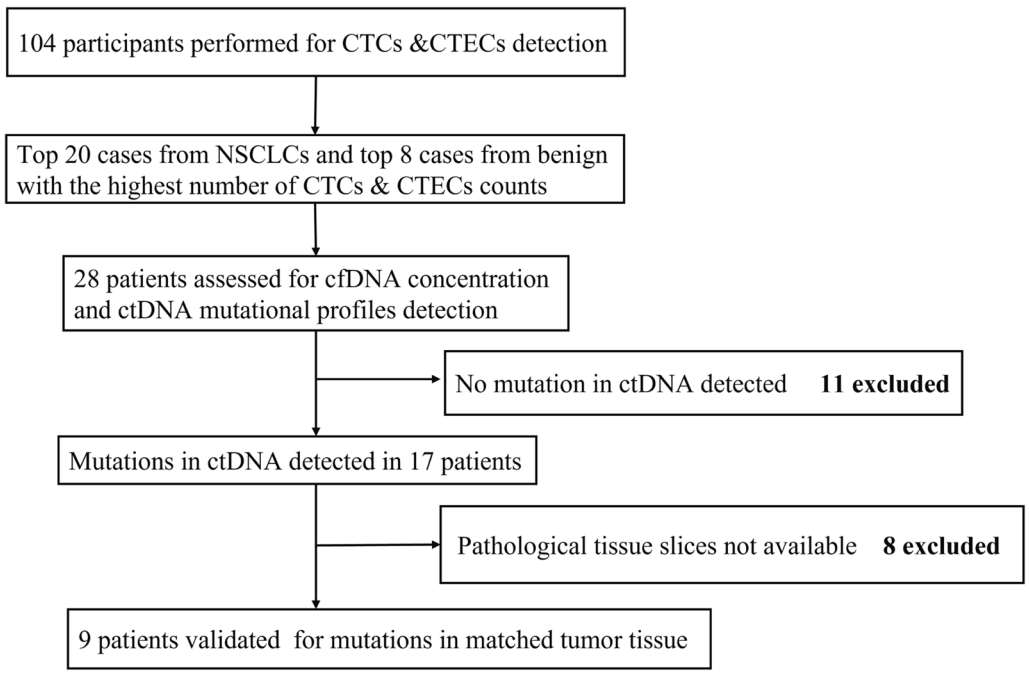


Fig. 1 Flow chart detailing the inclusion and exclusion procedures

5–10 pieces. After extraction, the quality of DNA was evaluated by Nanodrop One (Thermo Fisher, CA, USA), and the extracted DNA was quantified by Qubit 3.0 using the dsDNA HS Assay Kit (Life Technologies, CA, USA), according to the manufacturers' guidelines.

Detection and identification of CTCs and CTECs

Subtraction enrichment and immunostaining-fluorescence in situ hybridization (SE-iFISH) were applied to CTCs and CTECs detection. Firstly, target cells were enriched from 6 ml peripheral blood after the process of removing the white blood cells by centrifugation and incubation with immuno-magnetic beads, according to the manufacturer's protocols (Cytelligen, CA, USA). Secondly, the enriched cells were stained with fluorescent-labeled antibodies, including leukocyte marker (Alexa Fluor 594-CD45, red), endothelial marker (Alexa Fluor 488/Cy5-CD31, green/gold), tumor epithelial marker (Alexa Fluor 488/Cy5-CK18, green/gold), tumor immune marker (Alexa Fluor 488-PD-L1, green), and mesenchymal marker (Cy7-Vimentin, cyan). Next, after in situ hybridization using chromosome 8 centromere (CEP8) probe (Orange), target cells were initially screened and counted by the Metafer-Image Scanning System (Carl Zeiss, Oberkochen, Germany; MetaSystems, Altlussheim, Germany; Cytelligen). After manual review, CTCs and CTECs were finally identified. All the procedures and the standard morphological criteria were detailed in our previously published article [33]. Notably, vimentin-positive cells, which were not stained in all patients, were not included in the final counts.

ctDNA library preparation and NGS

At first, EDTA-anticoagulated blood samples were centrifuged at room temperature 200 g for 12 min, the supernatant was taken and centrifuged at room temperature 1200 g for 5 min to remove platelets and then centrifuged again at 4 °C for 16000 g for 10 min. The prepared plasma was frozen at -80 °C until further NGS library construction. Isolating cfDNA from 4 ml plasma using the AVENIO cfDNA Isolation Kit according to the manufacturer's instructions (Roche, Mannheim, Germany). After that, the plasma-extracted cfDNA was quantified using the Qubit dsDNA HS Assay Kit with Qubit 3.0 fluorimeter (ThermoFisher Scientific, CA, USA) and qualified using High Sensitivity DNA Chip with Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA). Next, the NGS libraries were prepared with DNA input of 50 ng using the AVENIO ctDNA Library Prep kit, AVENIO ctDNA Enrichment kit, AVENIO ctDNA Expanded Panel (targeted total of 77 genes, see Additional file 3: Table S5), and AVENIO Post-Hybridization kit. After quality control (QC) of the enriched libraries, the samples were performed for pooling and sequencing with NextSeq 500/550 High Output Kit v2 (300 cycles) on the Illumina NextSeq 500 (Illumina, CA, USA).

Tumor DNA library preparation and NGS

All samples were sequenced in a Clinical Laboratory Improvement Amendments (CLIA)- and College of American Pathologists (CAP)-certified genomic testing facility (Geneseeq Technology Inc., Nanjing, China). After being isolated from the tumor tissues, DNA fragments underwent end-repairing, A-tailing, and ligation with indexed adapters, and were selected as the size of 200 bp using Agencourt AMPure XP beads. Sequencing libraries were prepared by using the KAPA Hyper DNA Library Prep Kit (KAPA Biosystems, Wilmington, MA) according to the manufacturer's protocol. Hybridization-based target enrichment was performed with customized xGen lockdown probes (Integrated DNA Technologies) targeting 474 pan-solid tumor-relevant genes (Geneseeq radiotron®, Geneseeq Technology Inc.), the list of 474 genes, see Additional file3: Table S6. Captured libraries were PCR-amplified with KAPA HiFi Hot Start Ready Mix (KAPA Biosystems Wilmington, MA) followed by quantification using the KAPA Library Quantification kit (KAPA Biosystems Wilmington, MA, USA). DNA sequencing was performed on the HiSeq4000 NGS platform (Illumina) with a paired-end 150-bp read length.

WBCs DNA library preparation and NGS

Initially, WBCs DNA was extracted using AmoyDx Blood DNA Isolation Kit according to the manufacturer's instructions (Amoy Diagnostics Ltd, Xiamen, China). Subsequently, the DNA derived from the WBCs was assessed for quantification using the Qubit dsDNA HS Assay Kit (ThermoFisher). Sequencing libraries were prepared by using the Homologous Recombination Repair (HRR) Related 32 Gene Library Prep Kit (Amoy) and Genetic 61 Gene Library Prep Kit (Amoy) respectively, according to the manufacturer's protocol. The gene list of 32 Gene and 61 Gene panel were detailed in the Additional file3: Table S9, S10. After QC of the enriched libraries, the samples were performed for pooling and sequencing on the MiseqDx platform (Illumina) with a paired-end 150-bp read length.

Bioinformatic analysis

NGS analysis pipelines were prepared for data from blood samples and corresponding tumor tissues, respectively. The FASTQ files of ctDNA were analyzed using the AVENIO ctDNA Analysis Software, specifically version 1.1.0. For the analysis of WBCs DNA, the AmoyDx gHRR Analysis Software (version 1.5.0) and AmoyDx 61Gene Analysis Software (version 0.6.2) were employed for the

whole process. For tumor tissues, Trimmomatic (version 0.36) was initially performed for the FASTQ files quality control. After that, high-quality reads were aligned to the reference human genome (hg19, GRCh37) through Burrows-Wheeler aligner (BWA) v0.7.12. Deduplication was removed with Picard, and local realignment around indels and base quality score recalibration was performed with the Genome Analysis Toolkit (GATKv3.2). Furthermore, somatic single nucleotide variants (SNVs) and short insertions/deletions (indels) were identified by VarScan2, and copy number variations (CNVs) were detected using CNV Kit. CNV gain and loss were identified if the depth rate was ≥ 2.0 or ≤ 0.6 for the tissue sample. SNVs and indels were further filtered using the following criteria: (1) minimum ≥ 4 variant supporting reads and $\geq 2\%$ variant allele frequency (VAF) supporting the variant, (2) filtered if present in $> 1\%$ population frequency in the 1000 g or ExAC database, (3) filtered through an internally collected list of recurrent sequencing errors on the same sequencing platform. The final list of mutations was annotated using vcf2maf. The mean effective coverage depth was $> 3000\times$ for the plasma samples and $> 1000\times$ for tumor tissues. All the somatic variants were further classified based on NCCN/AMP/ASCO/CAP guidelines.

Meta-analysis

Study design: Case–control or cohort research articles published between 2019 and 2023 were included in the PubMed databases. The search terms used are shown as follows: non-small cell lung cancer, circulating tumor DNA, next generation sequencing, and tissue. The language was limited to English. The search strategy is designed as follows: (((non-small cell lung cancer[Title/Abstract]) AND (circulating tumor DNA[Title/Abstract])) AND ((next generation sequencing[Title/Abstract])) AND (tissue[Title/Abstract]))). Inclusion criteria: Research subjects: NSCLC patients; Intervention factor: NGS genotyping on ctDNA samples; Comparator factors: paired-tissue samples for genotyping; Outcome indicators: whether the ctDNA mutation profile was consistent with that in tumor tissue DNA. Exclusion criteria: review or meta-analysis or case report; repeated publications; absence of data; the research content does not align with the expectations; journal of publication not belonged to JCR partition Q1 or Q2; unpaired tissue samples. The detailed process of publications screening is presented in Additional file 1: Figure S1. *Data extraction*: Two reviewers extracted data from all eligible research, including the first author's name, clinical stage, ctDNA method, comparator method, and targeted gene. The testing results for multiple genes including true positive (TP), false positive (FP), false negative (FN), and true negative (TN) were collected. Genomic alterations in tissue genotyping were considered the "gold standard". All included studies underwent quality assessment using the standardized instrument Quality Assessment of Diagnostic Accuracy Tests (QUADAS-2) and summarized in Additional file 2: Figure S2.

Statistics analysis

SPSS Statistics software (version 25.0) was performed for the data analysis. The nonparametric Mann-Whitey U test (two groups) and Kruskal–Wallis test (multiple groups) were used to compare the measurement data respectively. The inspection level was $\alpha = 0.05$. ROC analysis was constructed to determine the cut-off scores for selected indicators and compare the diagnostic efficacy of selected indicators for NSCLC. Correlation analysis was performed using the nonparametric Pearson correlation. The inspection level $P < 0.05$ was defined as statistically significant. Statistical data was visualized using Prism 7.0 (GraphPad Software Inc., CA, USA) as well as OriginLab 2021 (Origin Software, MA, USA). For meta-analysis, the diagnostic data of TP, FP, FN, and TN were tabulated. The summary Receiver Operating Characteristic analysis was performed for meta-analysis of diagnostic accuracy. Software Review Manager 5.3 was used for relative analysis.

Results

Characterization of CTCs and CTECs in participants

A variety of CTCs and CTECs with different markers were observed. CK18 negative CTCs were detected in 97.9% (48/49), 83.9% (26/31), and 95.8% (23/24) in the NSCLC, benign lung disease, and healthy control, respectively. Remarkably, CK18-positive CTCs were detected in two NSCLC patients, with a tumor marker-positive detection rate of 4.2% (2/48). The circulating tumor microemboli (CTM) were detected in 4 lung cancer patients, while not in the other groups. Similarly, CTECs were detected in almost all participants. Vimentin-positive cells and CTM were detected in three squamous cell carcinama (SCC) patients. Several typical cells identified are photographed and listed, see Fig. 2. Compared with the benign group, the CTCs and CTECs counts are 5 units/6 ml (median, 1-11) and 5 units/6 ml (median, 2-7), respectively. The CTCs and CTECs counts of NSCLC patients were significantly increased, being 16 units/6 ml (median, 7–22.5) and 18 units/6 ml (median, 11.5–28.5), respectively (Additional file 3: Table S1; $P < 0.001$, $P < 0.001$, Fig. 3A-a, Fig. 3B-a). Additionally, the number of CTCs or CTECs was also higher in different sizes or different karyotypes of the lung cancer group compared to the benign group. (ALL: $P < 0.05$, Fig. 3A-b,c and B-b,c; Fig. 4A and B; Additional file 3: Table S2).

Diagnostic efficacy of CTCs and CTECs in NSCLC

To investigate the diagnostic efficacy of CTCs and CTECs in NSCLC, ROC analysis was performed based on multiple indicators of CTCs and CTECs using benign lung diseases as control. To illustrate the value of CTCs and CTECs for NSCLC, we selected three serum tumor markers (SCC, CEA, and CYFRA 21-1) as control, and any one of them higher than the cut-off value was considered positive. It was shown that the sensitivity and specificity of total CTCs and CTECs in the diagnosis of NSCLC were 67.3% and 77.6% [AUC (95%CI): 0.815 (0.722–0.907)], 83.9% and 77.4% [AUC (95%CI): 0.739 (0.618–0.860)], respectively, which were superior to serum tumor biomarkers or imaging evaluation (Tables 2, 3). Remarkably, small CTCs and triploid CTCs had high specificity in the diagnosis of NSCLC, which were 93.5% and 96.8% respectively. CTCs and CTECs with different sizes and karyotype have varying diagnostic sensitivity for different pathological types and stages of non-small cell lung cancer (detailed in Additional file 3: Table S3, S4). Furthermore, paired-wise combinations between CTCs and CTECs subgroups were conducted based on morphology and karyotype. After combination, the area under the ROC curves increased to Total(CTCs + CTECs) 0.826, ($_S$CTCs + $_S$CTECs) 0.898, and triploid(CTCs + CTECs) 0.872, respectively, indicating that the combined indicators were more effective than the single one (Table 2, Table 3).

Cell-free DNA level in NSCLC patients with high CTCs and CTECs counts

The concentration of cell-free DNA was 7.64 ± 3.79 ng. ml^{-1} in NSCLC patients (n = 20), compared with 5.64 ± 1.67 ng.ml^{-1}in the benign group (n = 8) (Table 4). Interestedly, ctDNA mutations were detected in patients with a significantly higher concentration of cfDNA (Fig. 5A). Remarkably, cfDNA was significantly higher in the III-IV stage group as well as the squamous cell carcinoma group (Fig. 5B, C). Furthermore, we conducted a thorough analysis of the correlation between cfDNA concentration and counts of CTCs and CTECs, as well as the concentration of serum biomarkers. Our findings indicated no linear correlation between different groups, such as between cfDNA concentration and CTCs counts (r = 0.150, P = 0.447), cfDNA concentration and CTECs counts (r = 0.008, P = 0.966), cfDNA and SCC concentration (r = 0.055, P = 0.800), and cfDNA and CEA concentration (r = − 0.211, P = 0.323). However, we did observe a positive correlation between cfDNA level and tumor size or maximum tumor diameter (MTD) (r = 0.430, P = 0.022), as well as CYFRA 21–1 concentration (r = 0.411, P = 0.041), (Fig. 5E, F, G, H) .

Comparison of mutational profiles between ctDNA and tumor DNA

The ctDNA mutations were detected in 17 out of the 28 patients (60.71%), including 4 in the benign group (50.00%), 5 in the lung adenocarcinoma (AC) group (41.70%), and 8 in the SCC group (100%) (Fig. 6B). Among ctDNA mutations, the TP53 mutation was the most common mostly appearing in the SCC group (see Additional file 3: Table S7).

A total of 5 mutations were found in 4 benign patients, 3 of them belonged to Class III variants: *RB1* (p.Arg621His), 48.9%; *ERBB2* (p.Lys937Arg), 0.34%; *MET* (p.Arg1005Gly), 0.09% and 2 were Class II mutations: *IDH1* (p.Arg132Ser), 0.05%; *PTEN* (p.Asn323fs), 0.12%. Since these 5 mutations were found in patients with benign lesions, we validated these mutations using WBCs DNA from corresponding patients and found *RB1* gene mutation (p.Arg621His, 48.9%) presented in WBCs DNA (see Additional file 3: Table S11). The remaining 4 mutations, varying from 0.05% to 0.34%, were not detected in WBCs DNA and suggested they were somatic mutations. It is worth mentioning that the patient with *PTEN* (p.Asn323fs) was suspected to be malignant based on imaging examinations, even though his bronchial biopsy was diagnosed as benign.

In order to validate the mutations, found in ctDNA, also presented in tDNA, NGS was performed in matched tDNA from 9 patients (2 with benign lesion, 4 with adenocarcinoma, 3 with SCC). The panel for tDNA contained 474 pan-solid tumor-relevant genes, which covered 75 genes in the ctDNA panel except 2 genes *CCND2* and *CCND3* (Fig. 6C). The full gene

(See figure on next page.)

Fig. 2 Multi-Fluorescence circulating tumor cells (CTCs) and circulating tumor-derived endothelial cells (CTECs) in patients with lung cancer: single-tumor-biomarker SE-iFISH (five/six-channel): nucleus (blue), CD45 (red), CD31 (green/yellow), CK18 (green/yellow), vimentin (cyan) and CEP8 (orange). **A** CTCs and CTECs with multiple markers in different sizes. **A-a** DAPI+/CD45-/CK18-/CD31-/VIM-/CEP8 small triploid CTC with adjacent WBCs (red arrows). **A-b** DAPI+/CD45-/CD31-/CK18+/CEP8 small diploid CTC. **A-c**.DAPI+/CD45-/CK18-/CD31-/VIM-/CEP8 large multiploid CTC with adjacent WBC (red arrow). **A-d** DAPI+/CD45-/CD31+/CK18-/CEP8 large multiploid CTEC. **A-e** DAPI+/CD45-/CD31+/CK18+/CEP8 small multiploid CTEC **B** Vimentin+ Mesenchymal cells from patients with lung squamous cell carcinoma. **B-a** DAPI+/CD45-/CK18-/CD31-/VIM+/CEP8 diploid mesenchymal cell. **B-b** DAPI+/CD45-/CK18-/CD31+/VIM+/CEP8 diploid mesenchymal cell. **C.** Circulating tumor microemboli (CTM) from patients with squamous cell carcinoma with adjacent WBCs (red arrows). Bars: 5μm

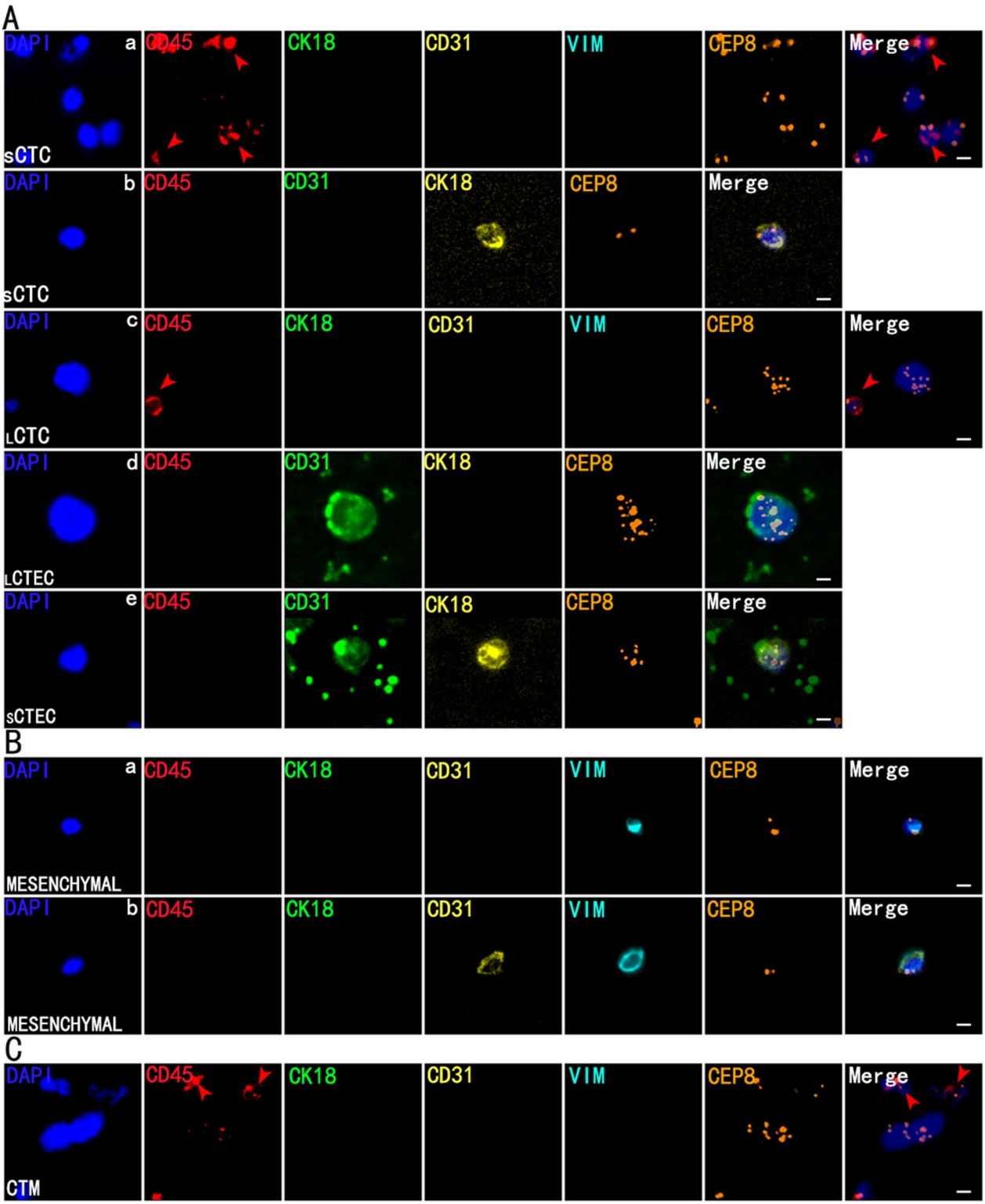


Fig. 2 (See legend on previous page.)

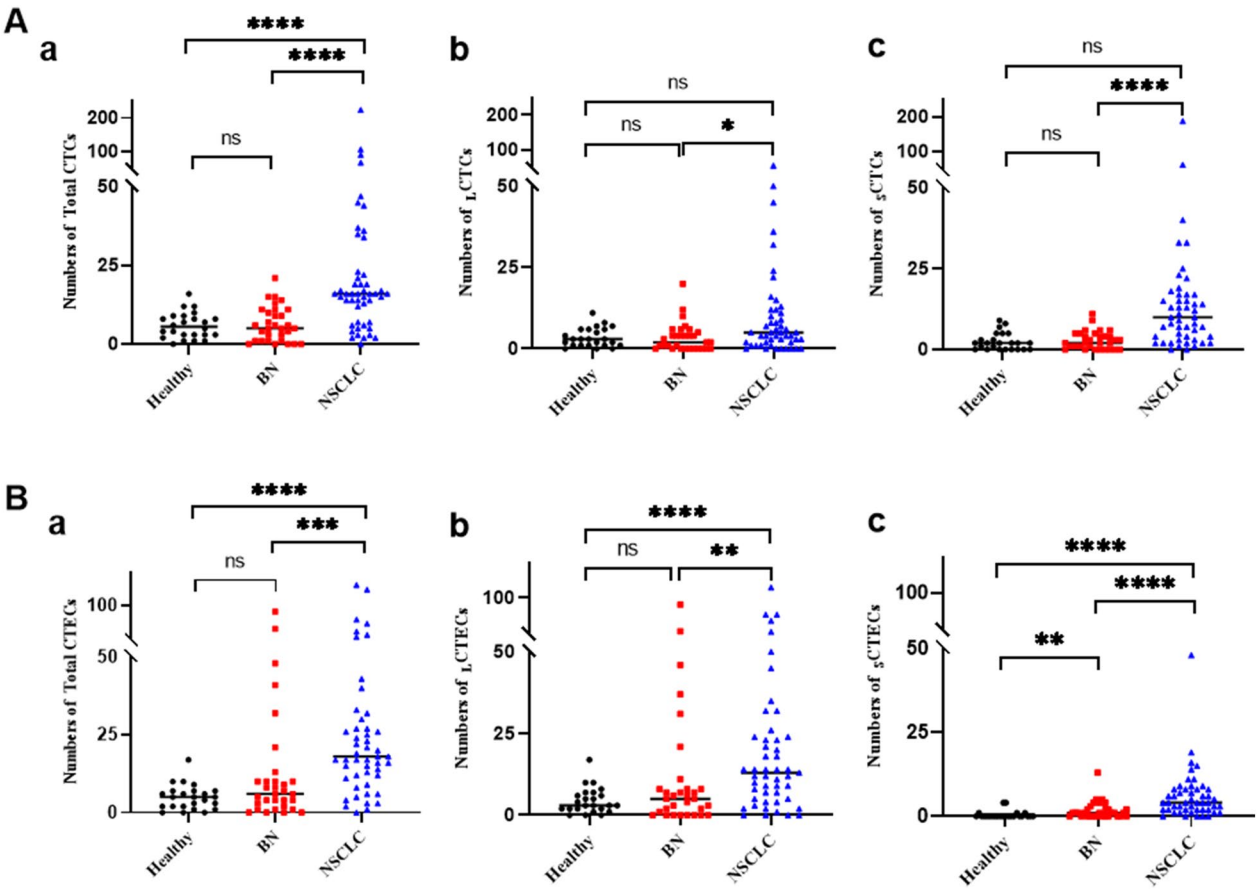


Fig. 3 Comparison of CTCs and CTECs counts among healthy, benign lung disease (BN) and NSCLC group. **A-a** Comparison of total CTCs counts among the groups. **A-b** Comparison of large CTCs ($_L$CTCs) counts among the groups. **A-c** Comparison of small CTCs ($_S$CTCs) counts among the groups. **B-a** Comparison of total CTECs counts among the groups. **B-b** Comparison of large CTECs ($_L$CTECs) counts among the groups. **B-c** Comparison of small CTECs ($_S$CTECs) counts among the groups. Data are presented as the median, $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, Mann–Whitney *U* test

list for both panels is shown in Additional file 3: Table S5 and Table S6. In these 9 patients, a total of 92 tDNA mutations were detected and 26 of them occurred in overlapping genes (see Additional file 3: Table S8). In detail, 1 mutation in 1 case of the benign lesion (another case of the benign lesion had no mutation detected), 11 mutations in 3 cases of AC, and 14 mutations in 3 cases of SSC (detailed mutation and frequency shown in Fig. 6D). The most common mutation was *EGFR* p.L858R and it was observed in 4 patients with AC and in 1 patient with SCC, but it was not detected in cfDNA. The second common mutation was CNV involved in *CDK4*. It was surprising that only 4 mutations *RB1* (p.R621H), *RET* (p.D631N), *P53* (p.V216L), and *TSC2* (p.R718H) were detected in both the tDNA and ctDNA. It suggested that the a lower coincidence among ctDNA and tDNA mutations. However, *PTCH1* (c.1347 + 6G > A) and *ALK* (Ser691Ser), which had a high frequency of 45.2% and 42.63% in plasma respectively, were not detected in the matched tumor tissue (Fig. 6D). Another interesting point is that tumor mutation burden (TMB) was higher in the SCC group and microsatellite instability-high (MSI-H) status was also observed in 1 SCC patient.

Since the sample size in our study is relatively small, we performed an additional meta-analysis from 21 articles containing ctDNA and tDNA sequencing data. The sensitivity, specificity, staging as well as methodology of each article are shown in Table 5. The rate of mutation to be detected or the sensitivity of mutation detection was different depending on the clinic staging. Among several studies involving advanced patients, the sensitivity ranged from 50 to 100% [34–53], However, the sensitivity was as low as 37% in studies focusing on early-stage patients [54] (Fig. 7A, Table 5). Moreover, the size of the panel used for detection also had an impact on the sensitivity. In two studies [40, 46], where only the *EGFR* or *KRAS* gene was covered, the sensitivity was significantly

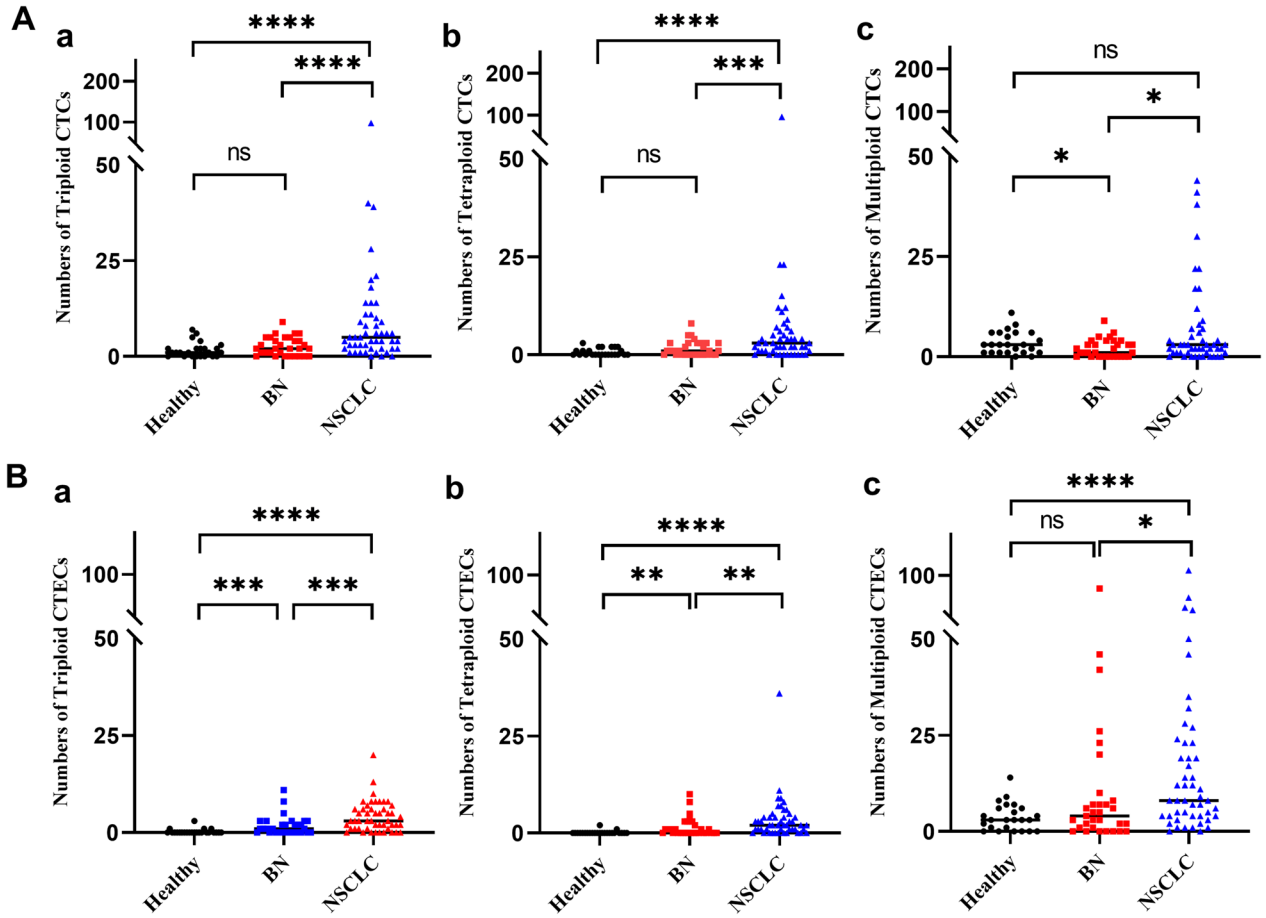


Fig. 4 Comparison of CTCs and CTECs counts in different aneuploids among healthy, benign lung disease (BN) and NSCLC groups. **A-a** Comparison of triploid CTCs counts among the groups. **A-b** Comparison of tetraploid CTCs counts among the groups. **A-c** Comparison of multiploid CTCs counts among the groups. **B-a** Comparison of triploid CTECs counts among the groups. **B-b** Comparison of tetraploid CTECs counts among the groups. **B-c** Comparison of multiploid CTECs counts among the groups. Data are presented as the median, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$, Mann–Whitney *U* test

higher compared to the first study in the list [54], which used a broader panel of genes (Table 5).

Discussion

Our study's findings indicated that CTCs and CTECs are valuable in aiding the diagnosis of lung cancer. Furthermore, our recently published study provides more comprehensive results, including patients with solitary pulmonary nodules, which further confirms the diagnostic value of this liquid biopsy method for early-stage lung cancer [33]. Therefore, in this article, we mainly discuss the role of ctDNA detection for early diagnosis and the coincidence between ctDNA mutation and tDNA mutation in order to suitably use it in clinical.

It is well known that cfDNA can be rapidly cleared in circulation, and it is difficult to capture due to low concentration and a very short half-life of less than one hour. Additionally, the proportion of ctDNA in cfDNA undergoes a profound alteration as cancer advances, ranging from 0.1 to 90% [55]. Thus, high sensitivity and stable assays are important for both cfDNA isolation and ctDNA mutation profiling. Currently, multiple methods have been applied for ctDNA mutation detection, such as ddPCR, BEAMing, Tagged-Amplicon deep sequencing (TAm-seq), Cancer Personalized Profiling by deep sequencing (CAPP-Seq), and Whole-genome-sequencing (WGS) or Whole-exome sequencing (WES). Compared to PCR-based assays, which can only detect a limited number of mutations, NGS-based methods utilizing unique molecular barcodes can identify and quantify multiple target genes simultaneously. This allows for the detection of mutant allele fractions (MAF) as low as < 0.1% [56]. In our research, we applied the Roche AVENIO ctDNA Expanded Kit and Illumina NextSeq 500 platform to detect ctDNA mutations in targeted 77 genes. Since low-abundance ctDNA mutations, such as

Table 2 Statistical parameters for ROC analysis of CTCs and CTECs for NSCLC

Test item	AUC	Cutoff value	Std. Error	P	95% CI
Total CTCs	0.815	11.5	0.047	< 0.0001	0.722–0.907
$_{L}$CTCs	0.663	6.5	0.061	0.015	0.544–0.782
$_{S}$CTCs	0.815	6.5	0.046	< 0.0001	0.724–0.906
Triploid CTCs	0.809	6.5	0.047	< 0.0001	0.717–0.901
Tetraploid CTCs	0.698	4.5	0.056	< 0.001	0.634–0.852
Multiploid CTCs	0.646	6.5	0.062	0.029	0.525–0.766
Total CTECs	0.739	10.5	0.062	< 0.001	0.618–0.860
$_{L}$CTECs	0.691	8.5	0.064	0.004	0.566–0.815
$_{S}$CTECs	0.781	2.5	0.053	< 0.0001	0.679–0.884
Triploid CTECs	0.716	3.5	0.058	0.001	0.602–0.830
Tetraploid CTECs	0.677	2.5	0.063	0.008	0.555–0.799
Multiploid CTECs	0.675	7.5	0.063	0.009	0.550–0.799
Total CTCs + CTECs	0.826	–	0.047	< 0.0001	0.734–0.917
$_{S}$CTCs + $_{S}$CTECs	0.898	–	0.036	< 0.0001	0.828–0.968
Triploid(CTCs + CTECs)	0.872	–	0.040	< 0.0001	0.794–0.950

Large CTCs, $_{L}$CTCs, small CTCs, $_{S}$CTCs, Large CTECs, $_{L}$CTECs, small CTECs, $_{S}$CTECs

Table 3 Evaluation of the diagnostic efficiency of single and combined tests for NSCLC [%]

Test item	SEN%	SPE%	PPV%	NPV%	AC%	LR +
Total CTCs	67.3	83.9	86.8	61.9	73.8	4.2
$_{L}$CTCs	38.8	67.7	65.5	41.2	50.0	1.2
$_{S}$CTCs	65.3	93.5	94.1	63.0	76.3	10.1
Triploid CTCs	57.1	96.8	96.6	58.8	72.5	17.7
Tetraploid CTCs	51.0	87.1	86.2	52.9	65.0	4.0
Multiploid CTCs	28.6	96.8	93.3	46.2	55.0	8.9
Total CTECs	77.6	77.4	84.4	68.6	77.5	3.4
$_{L}$CTECs	67.3	67.7	76.7	56.8	67.5	2.1
$_{S}$CTECs	67.3	77.4	82.5	60.0	71.3	3.0
Triploid CTECs	49.0	90.3	88.9	52.8	65.0	5.1
Tetraploid CTECs	36.7	71.0	66.7	41.5	50.0	1.3
Multiploid CTECs	59.2	74.2	78.4	53.5	65.0	2.3
SCC + CEA + CYFRA 21–1	52.2	75.8	77.4	50.0	61.3	2.17
CT/PET-CT diagnosis	57.1	77.4	80.0	53.3	65.0	2.48

Large CTCs, $_{L}$CTCs; small CTCs, $_{S}$CTCs; Large CTECs, $_{L}$CTECs; small CTECs, $_{S}$CTECs

SEN Sensitivity; *SPE* Specificity; *PPV* Positive predictive value; *NPV* Negative predictive value; LR+ Positive likelihood ratio; *SCC* Squamous cell carcinoma antigen; *CEA* Carcinoma embryonic antigen; *CYFRA21-1* Cytokeratin fragment antigen 21–1

ERBB2 and *IDH1* with the percentage of variants 0.34% and 0.05% respectively, had been caught in one patient with the benign lesion, which suggested that this kit is relatively sensitive and NGS-based ctDNA detection may have advantages in early finding mutant ctDNA.

In this study, even though the cell-free DNA concentration showed a significant positive association with the tumor size and stage, we still cannot confirm how much of this cfDNA is derived from the tumor tissue, since the cfDNA level can also increase in many instances such as inflammation, surgery or drug stimulation which could cause cfDNA release by normal cells or hematopoietic cells [32]. In addition, we found no correlation between cfDNA concentration and CTCs or CTECs counts, which does bolster the previous research that most cfDNA is not derived from circulating tumor cells [57].

One unanticipated finding in this research was the discordance in mutations between ctDNA from blood and DNA derived from matched tumor tissue. On the one hand, low-frequency variants such as *IDH1* (Arg132Gly),

Table 4 Statistics of the cfDNA concentration in enrolled 28 patients (ng/ml)

Group	No(%)	Mean	S.D	95%CI	Median	Range	P25-P75
Benign	8	5.64	1.67	4.11–8.47	5.16	4.36	4.24–7.28
NSCLC	20	7.64	3.79	4.79–9.38	7.29	13.71	4.42–9.81
Pathological type							
AC	12(60%)	6.10	2.83	4.30–7.89	4.95	7.71	3.95–9.11
SCC	8(40%)	9.94	4.03	6.57–13.31	8.62	10.16	6.24–13.93
TNM stage							
I-II	15(75%)	6.42	2.67	4.94–7.89	5.84	7.71	4.01–9.00
III-IV	5(25%)	11.29	4.59	5.59–16.99	13.73	9.87	6.36–15.00

NSCLC non-small cell lung cancer; *AC* Adenocarcinoma; *SCC* Squamous carcinoma

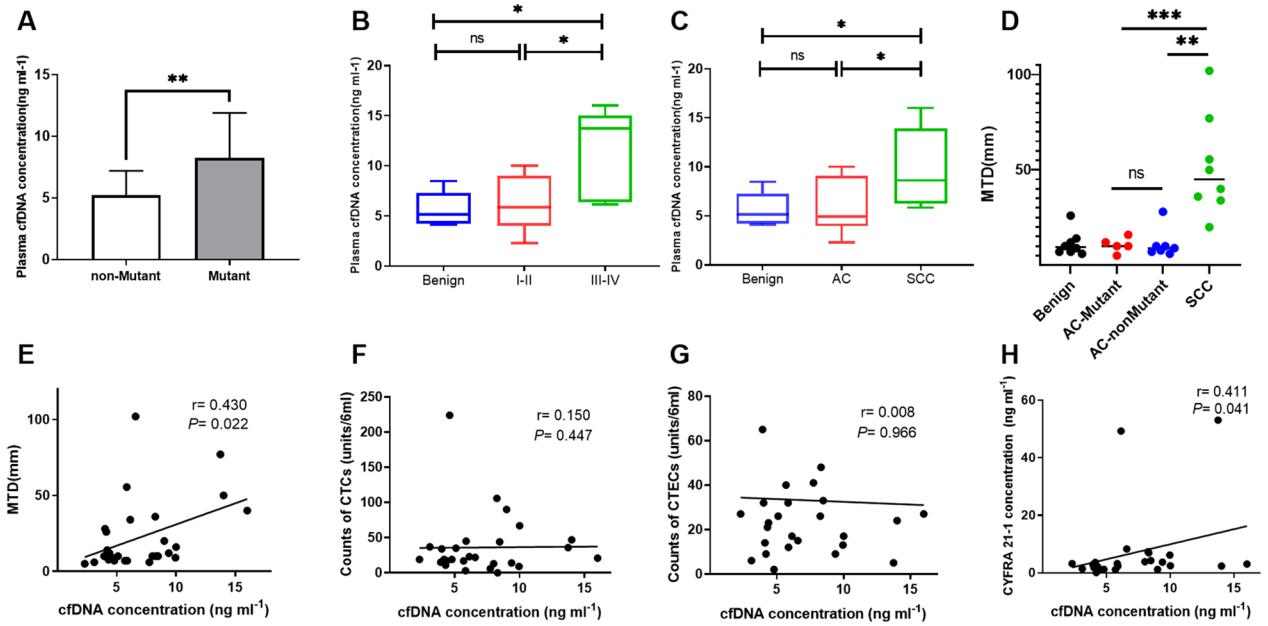


Fig. 5 Analysis of the plasma cfDNA concentration in enrolled 28 patients. **A** Plasma cfDNA concentration between patients with or without ctDNA mutations**. B** Plasma cfDNA concentration between different TNM stages. **C** Plasma cfDNA concentration between different pathological types. **D** Tumor size (Maximum tumor diameter, MTD) between different pathological types. **E** The correlation between cfDNA concentration and tumor size (MTD). **F** The correlation between cfDNA concentration and counts of CTCs. **G** The correlation between cfDNA concentration and counts of CTECs. **H** The correlation between the concentration of cfDNA and CYFRA 21–1.*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$, Mann–Whitney *U* test

TP53 (c.919 + 1G > A), *GNAS* (Arg201Cys) and *MSH* (Try1066Cys), as well as high-frequency variants *PTCH1*(c.1347 + 6G > A) and *ALK* (Ser691Ser) detected in the plasma of patients were not found in the matched tumor. We considered that those low-frequency variants may arise from minor subclones due to temporal and spatial heterogeneity of the tumor itself [58], while the high-frequency mutations may be related to clonal hematopoiesis (CH) [30, 59–61]. On the other hand, many mutations in tumor tissue especially for *EGFR* (L858R) have not been detected in the plasma. This may be possibly due to the minimal release of early-stage tumors. Unlike CTCs that are actively discharged from the tumor tissues and can be viewed as small metastases, known as cM0(i +) staging [62], ctDNA is released passively by apoptotic and necrotic cancer cells. Therefore, external factors like surgery, radiation, or chemotherapy have the potential to promote the release of tumor DNA into the bloodstream. Additionally, tumor heterogeneity may also be attributed to this concordance. Our meta-analysis results also show the sensitivity of ctDNA testing using NGS is higher in advanced NSCLC but not satisfactory in early-stage patients, which is similar to our findings. Thus, it means that a high-intensity ctDNA assay is needed to capture these rare mutations.

Table 5 Characteristics of included studies for meta-analysis

Study	Stage	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	ctDNA Method	tDNA Method	Targeted Gene
Bin Zhang et al	early-stage	15	5	25	3	0.38 (0.23–0.54)	0.38 (0.09–0.76)	NGS	NGS	Panel Targeted 546 GENES*
Charu Aggarwal et al	IV	54	11	21	42	0.72 (0.60–0.82)	0.79 (0.66–0.89)	NGS	NGS	Panel Targeted 73 GENES;153GENES*
Chenglong Zhao et al	Advanced	381	21	111	6	0.77 (0.73–0.81)	0.22 (0.09–0.42)	NGS	NGS	Panel Targeted 168 GENES
García-Pardo et al	Advanced	35	5	1	28	0.97 (0.85–1.00)	0.85 (0.68–0.95)	NGS	NGS	EGFR; ALK;ROS1;MET;ERBB2;BRAF;KRAS
Hao Bai et al	I-IV	20	11	6	42	0.77 (0.56–0.91)	0.79 (0.66–0.89)	NGS	AMRS-PCR	EGFR
Jian-xin Yin et al	IA-IV	7	5	0	64	1.00 (0.59–1.00)	0.93 (0.84–0.98)	NGS	NGS	EGFR; KRAS
Jun Cai et al	Advanced	30	3	16	0	0.65 (0.50–0.79)	0.00 (0.00–0.71)	NGS	NGS	Panel Targeted 95 GENES*
Justin Jee et al	Advanced	216	3	163	538	0.57 (0.52–0.62)	0.99 (0.98–1.00)	NGS	NGS	Panel Targeted 24GENES*
Lawrence Hsu Lin et al	II- IV	27	3	23	18	0.54 (0.39–0.68)	0.86 (0.64–0.97)	NGS	NGS	Panel Targeted 74 GENES*
Luis E. Raez et al	IV	14	18	7	114	0.67 (0.43–0.85)	0.86 (0.79–0.92)	NGS	NGS	Panel Targeted 86 GENES*
Maria Gabriela O. et al	II IV	51	3	12	61	0.81 (0.69–0.90)	0.95 (0.87–0.99)	NGS	NGS	Panel Targeted 24 GENES*
Martin Metzenmacher et al	IV	10	3	3	0	0.77 (0.46–0.95)	0.00 (0.00–0.71)	NGS	NGS	Panel Targeted 73 GENES*
Moom R. et al	I-IV	89	66	62	3	0.59 (0.51–0.67)	0.04 (0.00–0.12)	NGS	NGS	ALK,EGFR,ROS1,BRAF,RET,MET, ERBB2, KRAS
Paul van et al	Advanced	33	4	17	12	0.66 (0.51–0.79)	0.75 (0.48–0.93)	NGS	NGS	EGFR, KRAS, BRAF, ERBB2, and PIK3CA
Peng-Peng Kuang et al	IB-III	19	0	19	0	0.50 (0.33–0.67)	Not estimable	NGS	NGS	Panel Targeted 425 GENES*
Ramon Palmero et al	Advanced	29	17	19	3	0.60 (0.45–0.74)	0.15 (0.03–0.38)	NGS	NGS	ALK,EGFR,ROS1,BRAF,RET, MET, ERBB2, KRAS
Sehhoon Park et al	Advanced	108	11	53	87	0.67 (0.59–0.74)	0.89 (0.81–0.94)	NGS	NGS	Panel Targeted 74 GENES*
Sheehyun Kim et al	Advanced	28	18	19	0	0.60 (0.44–0.74)	0.00 (0.00–0.19)	NGS	NGS	Panel Targeted 171 and 25 fusion genes*
Vassiliki A. et al	Advanced	566	53	287	754	0.66 (0.63–0.70)	0.93 (0.91–0.95)	NGS	PCR	EGFR
Yukti Choudhury et al	I-IV	9	2	3	10	0.75 (0.43–0.95)	0.83 (0.52–0.98)	NGS	NGS	Panel Targeted 49 GENES*
Zeyun Lin et al	I-IV	81	6	34	69	0.70 (0.61–0.79)	0.92 (0.83–0.97)	NGS	NGS	Panel Targeted 168 GENES*

EGFR epidermal growth factor receptor; *ALK* anaplastic lymphoma kinase; *ROS1 ROS* proto-oncogene 1, receptor tyrosine kinase; *HER2/ERBB2* human epidermal growth factor receptor 2; *BRAF* serine/threonine-protein kinase B-Raf; *MET* MET proto-oncogene, receptor tyrosine kinase; *RET RET* proto-oncogene; *KRAS* proto-oncogenes, *K-rasFN*, false negative; *FP* false positive; *TP* true positive; *TN* true negative; *NGS* next generation sequencing

* The genes panels some studies applied include 8 core driver genes in lung cancer: EGFR、ALK、ROS1、BRAF、RET、MET、HER2、KRAS

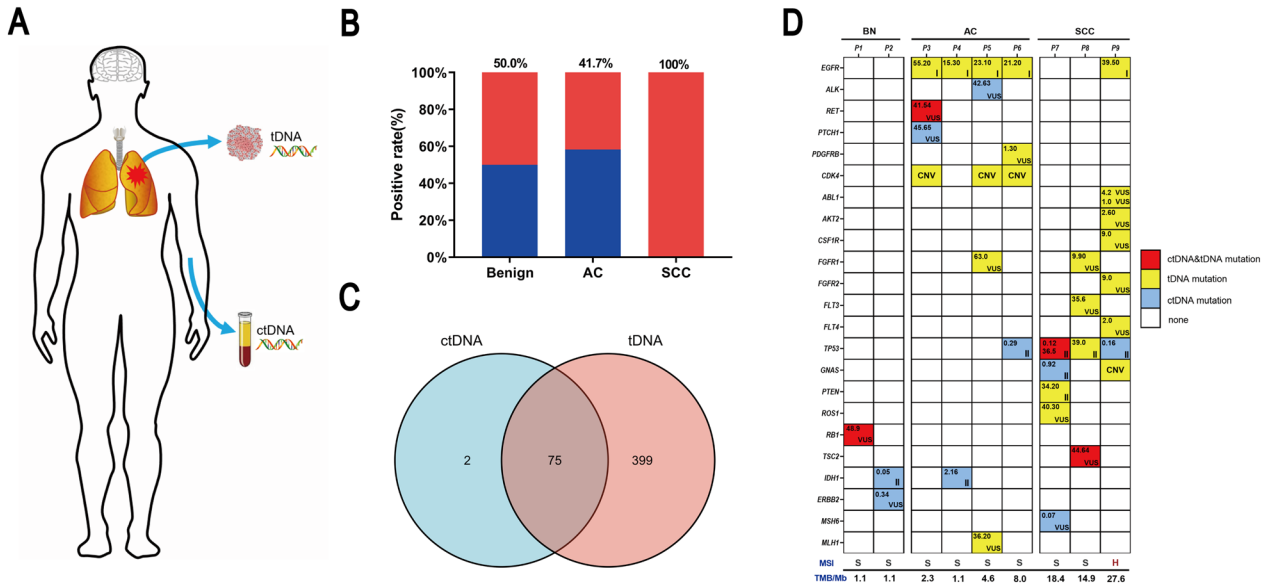


Fig. 6 Characteristics of the mutant genes detected in circulating tumor DNA (ctDNA) and tumor DNA (tDNA). **A** Schematic figure illustrating the sources of ctDNA and tDNA. **B** The positive rates of ctDNA mutations among benign, adenocarcinoma (AC), and squamous cell carcinoma (SCC) subgroups of the enrolled 28 patients. **C** Venn figure showing the target genes covered in panels of both ctDNA and tDNA. **D** Heatmap detailing the mutant genes detected in the ctDNA and matched tDNA of the 9 patients with fractional abundance in the upper left corner as well as the microsatellite instability (MSI) status and tumor mutation burden (TMB) score; VUS: variant of uncertain significance

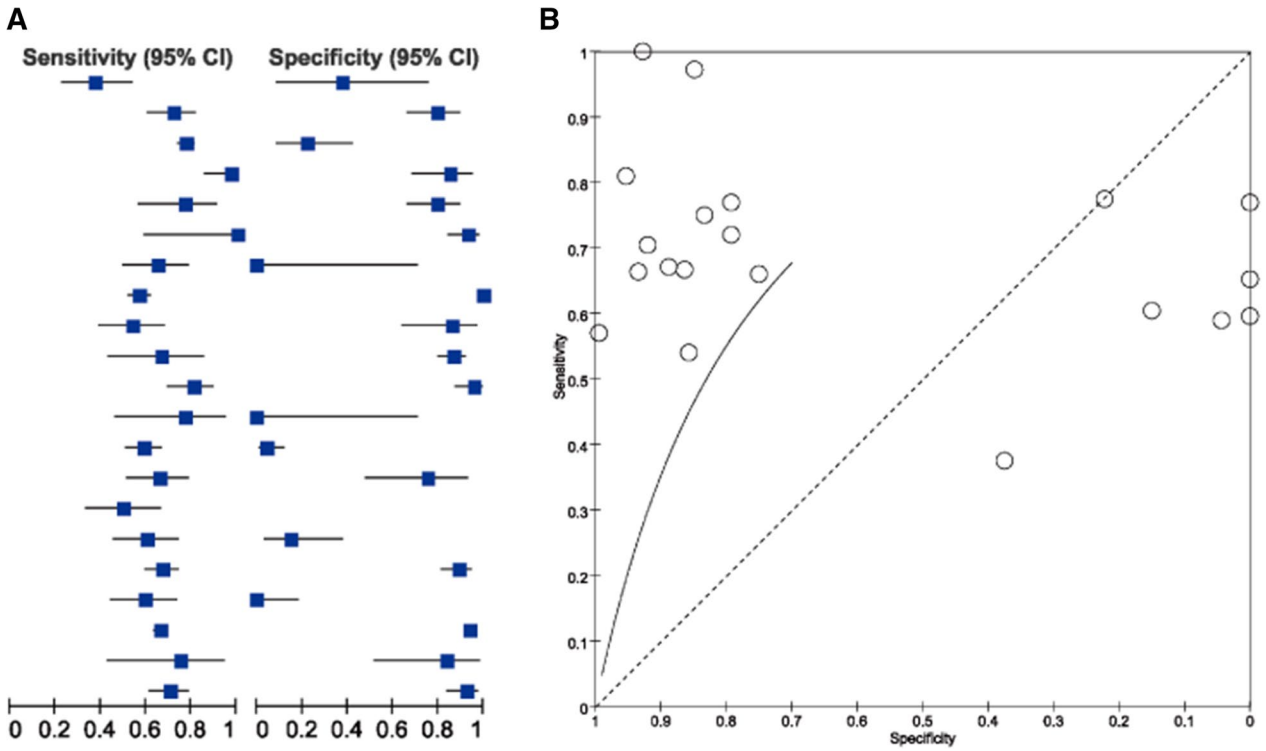


Fig. 7 Meta-Analysis of results for the ctDNA genotyping vs tissue genotyping of 21 studies. **A** Forest plot of the analysis for the ctDNA genotyping. **B** sROC analysis for the ctDNA genotyping

Conclusions

In summary, the detection of CTCs and CTECs based on SE-iFISH has shown promising roles for NSCLC diagnosis. However, ctDNA detection provides an option for screening potential patients with cancer and monitoring treatment response. Despite liquid biopsy shows currently various dilemmas in clinical practice, with the progress of technology, the reduction of sequencing costs as well as big data analysis based on multiple indicators, it will eventually play a more important role in the diagnosis, treatment, or prognosis of various cancers in the future.

AC	Adenocarcinoma
CTCs	Circulating tumor cells
CTECs	Circulating tumor-derived endothelial cells
ctDNA	Circulating tumor DNA
NGS	Next generation sequencing
NSCLC	Non-small cell lung cancer
SCC	Squamous cell carcinoma
SE-iFISH	Subtraction enrichment and immunostaining-fluorescence in situ hybridization
tDNA	Tumor DNA

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-023-04746-8.

Additional file 1: Fig S1. Flowchart of publications selection in meta-analysis.

Additional file 2: Fig S2. Included studies quality assessment according to QUADAS-2. **A.** Risk of bias and applicability concerns summary: review authors' judgments about each domain for each included study **B.** Risk of bias and applicability concerns graph.

Additional file 3: Supplementary **Table 1.** Statistical parameters of CTCs and CTECs of 104 participating subjects; **Table 2.** Statistical parameters of aneuploid CTCs and CTECs of 104 participating subjects; **Table 3.** The diagnostic sensitivity of CTCs and CTECs in different pathological types and stages of NSCLC; **Table 4.** The diagnostic sensitivity of aneuploid CTCs and CTECs in different pathological types and stages of NSCLC; **Table 5.** Target gene information of ctDNA panel; **Table 6.** Target gene information of tDNA panel; **Table 7.** Mutational profile of ctDNA in enrolled 28 patients; **Table 8.** Mutational profiles of tumor DNA in filtered 9 patients; **Table 9.** Target gene information of HRR panel; **Table 10.** Target gene information of 61 GENE panel; **Table 11.** Mutational profiles of WBCs DNA in benign 4 patients.

Acknowledgements

We wish to express our sincere gratitude for the generous assistance from Dr. Peter. P. Lin. and his team (Cytelligen, CA, USA). We are also grateful for the kind support received from the team members of the Epione Medical Laboratory in Shanghai, China, the Clinical Laboratory at Nanjing Geneseeq Technology Inc. in Nanjing, China, and the Department of Prenatal Diagnosis Center at Shanghai General Hospital.

Author contributions

JC and QH contributed to the conception of the study and manuscript revision. JC, JX, BH, and YG performed the experiments. JX, BH, and SH participated in data analysis and manuscript preparation. JL assisted with pathologic diagnosis.

Funding

This study was financially supported by the Shanghai "Rising Stars of Medical Talents" Youth Development Program (No. SHWRS (2021)_099 to JC).

Availability of data and materials

All data generated during this study are available from the corresponding author by request.

Declarations

Ethics approval and consent to participate

The study was approved by the Ethical Review Board of Shanghai General Hospital, Shanghai Jiaotong University School of Medicine, China (2016KY130).

Consent for publication

All authors have approved the manuscript for submission.

Competing interests

All the authors declare that they have no competing interests.

Author details

[1]Cancer Center, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. [2]Department of Pathology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

Received: 10 July 2023 Accepted: 21 November 2023
Published online: 01 December 2023

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA A Cancer J Clin. 2021;71:209–49.
2. Feng R-M, Zong Y-N, Cao S-M, Xu R-H. Current cancer situation in China: good or bad news from the 2018 global cancer statistics? Cancer Commun. 2019;39:22.
3. Planchard D, Popat S, Kerr K, Novello S, Smit EF, Faivre-Finn C, et al. Metastatic non-small cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2018;29:iv192-237.
4. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. N Engl J Med. 2020;382:503–13.
5. Trimothy C, William BC, Denise RA, Christine DB, Kathy LC, Fenghai D, et al. Results of initial low-dose computed tomographic screening for lung cancer. New Eng J Med. 2013;368:1980.
6. Hofman P. Liquid biopsy for lung cancer screening: usefulness of circulating tumor cells and other circulating blood biomarkers. Cancer Cytopathol. 2021;129:341–6.
7. Qvick A, Stenmark B, Carlsson J, Isaksson J, Karlsson C, Helenius G. Liquid biopsy as an option for predictive testing and prognosis in patients with lung cancer. Mol Med. 2021;27:68.
8. Di Capua D, Bracken-Clarke D, Ronan K, Baird A-M, Finn S. The liquid biopsy for lung cancer: state of the art, limitations and future developments. Cancers. 2021;13:3923.
9. Freitas C, Sousa C, Machado F, Serino M, Santos V, Cruz-Martins N, et al. The role of liquid biopsy in early diagnosis of lung cancer. Front Oncol. 2021;11: 634316.
10. Kalluri R. EMT: when epithelial cells decide to become mesenchymal-like cells. J Clin Invest. 2009;119:1417–9.
11. Platel V, Faure S, Corre I, Clere N. Endothelial-to-mesenchymal transition (EndoMT): roles in tumorigenesis, metastatic extravasation and therapy resistance. J Oncol. 2019;2019:8361945.
12. Shi J, Zhao C, Shen M, Chen Z, Liu J, Zhang S, et al. Combination of microfluidic chips and biosensing for the enrichment of circulating tumor cells. Biosens Bioelectron. 2022;202: 114025.
13. Myung JH, Hong S. Microfluidic devices to enrich and isolate circulating tumor cells. Lab Chip. 2015;15:4500–11.
14. Grover PK, Cummins AG, Price TJ, Roberts-Thomson IC, Hardingham JE. Circulating tumour cells: the evolving concept and the inadequacy of their enrichment by EpCAM-based methodology for basic and clinical cancer research. Ann Oncol. 2014;25:1506–16.

15. Chen X, Zhou F, Li X, Yang G, Zhang L, Ren S, et al. Folate receptor-positive circulating tumor cell detected by LT-PCR-based method as a diagnostic biomarker for non-small-cell lung cancer. J Thorac Oncol. 2015;10:1163–71.
16. Li J, Liao Y, Ran Y, Wang G, Wu W, Qiu Y, et al. Evaluation of sensitivity and specificity of CanPatrol technology for detection of circulating tumor cells in patients with non-small cell lung cancer. BMC Pulm Med. 2020;20:274.
17. Mikolajczyk SD, Millar LS, Tsinberg P, Coutts SM, Zomorrodi M, Pham T, et al. Detection of EpCAM-negative and cytokeratin-negative circulating tumor cells in peripheral blood. J Oncol. 2011;2011: 252361.
18. Su Z, Wang Z, Ni X, Duan J, Gao Y, Zhuo M, et al. Inferring the evolution and progression of small-cell lung cancer by single-cell sequencing of circulating tumor cells. Clin Cancer Res. 2019;25:5049–60.
19. Lin PP. Aneuploid circulating tumor-derived endothelial cell (CTEC): a novel versatile player in tumor neovascularization and cancer metastasis. Cells. 2020;9:1539.
20. Aucamp J, Bronkhorst AJ, Badenhorst CPS, Pretorius PJ. The diverse origins of circulating cell-free DNA in the human body: a critical re-evaluation of the literature. Biol Rev Camb Philos Soc. 2018;93:1649–83.
21. Adalsteinsson VA, Ha G, Freeman SS, Choudhury AD, Stover DG, Parsons HA, et al. Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. Nat Commun. 2017;8:1324.
22. Jordan EJ, Kim HR, Arcila ME, Barron D, Chakravarty D, Gao J, et al. Prospective comprehensive molecular characterization of lung adenocarcinomas for efficient patient matching to approved and emerging therapies. Cancer Discov. 2017;7:596–609.
23. Goldberg SB, Narayan A, Kole AJ, Decker RH, Teysir J, Carriero NJ, et al. Early assessment of lung cancer immunotherapy response via circulating tumor DNA. Clin Cancer Res. 2018;24:1872–80.
24. US Food FDA and Drug Administration. Premarket approval P150044 Cobas EGFR MUTATION TEST V2. 2018. https://www.accessdata.fda.gov/scripts/ cdrh/cfdocs/cfpma/pma.cfm?id=P150044
25. Lindeman NI, Cagle PT, Aisner DL, Arcila ME, Beasley MB, Bernicker EH, et al. Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: guideline from the college of American pathologists, the international association for the study of lung cancer, and the association for molecular pathology. Arch Pathol Lab Med. 2018;142:321–46.
26. Yang Z, Qi W, Sun L, Zhou H, Zhou B, Hu Y. DNA methylation analysis of selected genes for the detectionof early-stage lung cancer using circulating cell-free DNA. Adv Clin Exp Med. 2018;28:355–60.
27. Zhang C, Yu W, Wang L, Zhao M, Guo Q, Lv S, et al. DNA methylation analysis of the SHOX2 and RASSF1A panel in bronchoalveolar lavage fluid for lung cancer diagnosis. J Cancer. 2017;8:3585–91.
28. Ponomaryova AA, Rykova EY, Cherdyntseva NV, Skvortsova TE, Dobrodeev AY, Zavyalov AA, et al. Potentialities of aberrantly methylated circulating DNA for diagnostics and post-treatment follow-up of lung cancer patients. Lung Cancer. 2013;81:397–403.
29. Moding EJ, Liu Y, Nabet BY, Chabon JJ, Chaudhuri AA, Hui AB, et al. Circulating tumor DNA dynamics predict benefit from consolidation immunotherapy in locally advanced non-small cell lung cancer. Nat Cancer. 2020;1:176–83.
30. Chae YK, Oh MS. Detection of minimal residual disease using ctDNA in lung cancer: current evidence and future directions. J Thorac Oncol. 2019;14:16–24.
31. Wan R, Wang Z, Lee JJ, Wang S, Li Q, Tang F, et al. Comprehensive analysis of the discordance of EGFR mutation status between tumor tissues and matched circulating tumor DNA in advanced non-small cell lung cancer. J Thorac Oncol. 2017;12:1376–87.
32. Chen M, Zhao H. Next-generation sequencing in liquid biopsy: cancer screening and early detection. Hum Genom. 2019;13:34.
33. Xie J, Ruan Z, Zheng J, Gong Y, Wang Y, Hu B, et al. Detection of circulating rare cells benefitted the diagnosis of malignant solitary pulmonary nodules. J Cancer Res Clin Oncol. 2022;148:2681–92.
34. García-Pardo M, Czarnecka-Kujawa K, Law JH, Salvarrey AM, Fernandes R, Fan ZJ, et al. Association of circulating tumor DNA testing before tissue diagnosis with time to treatment among patients with suspected advanced lung cancer: the accelerate nonrandomized clinical trial. JAMA Netw Open. 2023;6: e2325332.
35. Palmero R, Taus A, Viteri S, Majem M, Carcereny E, Garde-Noguera J, et al. Biomarker discovery and outcomes for comprehensive cell-free circulating tumor DNA versus standard-of-care tissue testing in advanced non-small-cell lung cancer. JCO Precis Oncol. 2021;5:93.
36. Kuang PP, Li N, Liu Z, Sun TY, Wang SQ, Hu J, Ou W, Wang SY, et al. Circulating tumor DNA analyses as a potential marker of recurrence and effectiveness of adjuvant chemotherapy for resected non-small-cell lung cancer. Front Oncol. 2021. https://doi.org/10.3389/fonc.2020.595650.
37. Fernandes MGO, Cruz-Martins N, Souto Moura C, Guimarães S, Pereira Reis J, Justino A, et al. Clinical application of next-generation sequencing of plasma cell-free DNA for genotyping untreated advanced non-small cell lung cancer. Cancers. 2021;13:2707.
38. Aggarwal C, Thompson JC, Black TA, Katz SI, Fan R, Yee SS, et al. Clinical implications of plasma-based genotyping with the delivery of personalized therapy in metastatic non-small cell lung cancer. JAMA Oncol. 2019;5:173–80.
39. Kim S, Kim S, Kim SH, Jung EH, Suh KJ, Kim YJ, et al. Clinical validity of oncogenic driver genes detected from circulating tumor DNA in the blood of lung cancer patients. Transl Lung Cancer Res. 2023;12:1185–96.
40. Yin JX, Hu WW, Gu H, Fang JM. Combined assay of circulating tumor DNA and protein biomarkers for early noninvasive detection and prognosis of non-small cell lung cancer. J Cancer. 2021;12:1258.
41. Metzenmacher M, Hegedüs B, Forster J, Schramm A, Horn PA, Klein CA, et al. Combined multimodal ctDNA analysis and radiological imaging for tumor surveillance in Non-small cell lung cancer. Transl Oncol. 2022;15: 101279.
42. Lin Z, Li Y, Tang S, Deng Q, Jiang J, Zhou C. Comparative analysis of genomic profiles between tissue-based and plasma-based next-generation sequencing in patients with non-small cell lung cancer. Lung Cancer. 2023;182: 107282.
43. Lin LH, Allison DHR, Feng Y, Jour G, Park K, Zhou F, et al. Comparison of solid tissue sequencing and liquid biopsy accuracy in identification of clinically relevant gene mutations and rearrangements in lung adenocarcinomas. Mod Pathol. 2021;34:2168–74.
44. Choudhury Y, Tan M-H, Shi JL, Tee A, Ngeow KC, Poh J, et al. Complementing tissue testing with plasma mutation profiling improves therapeutic decision-making for patients with lung cancer. Front Med. 2022;9: 758464.
45. Leest PVD, Janning M, Rifaela N, Azpurua MLA, Kropidlowski J, Loges S, et al. Detection and monitoring of tumor-derived mutations in circulating tumor DNA using the ultraSEEK lung panel on the massARRAY system in metastatic non-small cell lung cancer patients. IJMS. 2023;24:13390.
46. Bai H, Xia J, Zhao X, Gong Z, Zhang D, Xiong L. Detection of EGFR mutations using target capture sequencing in plasma of patients with non-small-cell lung cancer. J Clin Pathol. 2019;72:379–85.
47. Park S, Olsen S, Ku BM, Lee M-S, Jung H-A, Sun J-M, et al. High concordance of actionable genomic alterations identified between circulating tumor DNA-based and tissue-based next-generation sequencing testing in advanced non-small cell lung cancer: the Korean lung liquid versus invasive biopsy program. Cancer. 2021;127:3019–28.
48. Papadimitrakopoulou VA, Han J-Y, Ahn M-J, Ramalingam SS, Delmonte A, Hsia T-C, et al. Epidermal growth factor receptor mutation analysis in tissue and plasma from the AURA3 trial: osimertinib versus platinum-pemetrexed for T790M mutation-positive advanced non-small cell lung cancer. Cancer. 2020;126:373–80.
49. Raez LE, Brice K, Dumais K, Lopez-Cohen A, Wietecha D, Izquierdo PA, et al. Liquid biopsy versus tissue biopsy to determine front line therapy in metastatic non-small cell lung cancer (NSCLC). Clin Lung Cancer. 2023;24:120–9.
50. Jee J, Lebow ES, Yeh R, Das JP, Namakydoust A, Paik PK, et al. Overall survival with circulating tumor DNA-guided therapy in advanced non-small-cell lung cancer. Nat Med. 2022;28:2353–63.
51. Cai J, Jiang H, Li S, Yan X, Wang M, Li N, et al. The landscape of actionable genomic alterations by next-generation sequencing in tumor tissue versus circulating tumor DNA in Chinese patients with non-small cell lung cancer. Front Oncol. 2021;11: 751106.
52. Zhao C, Li J, Zhang Y, Han R, Wang Y, Li L, et al. The rational application of liquid biopsy based on next-generation sequencing in advanced non-small cell lung cancer. Cancer Med. 2023;12:5603–14.
53. Roosan MR, Mambetsariev I, Pharaon R, Fricke J, Husain H, Reckamp KL, et al. Usefulness of circulating tumor DNA in identifying somatic

mutations and tracking tumor evolution in patients with non-small cell lung cancer. Chest. 2021;160:1095–107.
54. Zhang B, Niu X, Zhang Q, Wang C, Liu B, Yue D, et al. Circulating tumor DNA detection is correlated to histologic types in patients with early-stage non-small-cell lung cancer. Lung Cancer. 2019;134:108–16.
55. Corcoran RB, Chabner BA. Application of cell-free DNA analysis to cancer treatment. N Engl J Med. 2018;379:1754–65.
56. Phallen J, Sausen M, Adleff V, Leal A, Hruban C, White J, et al. Direct detection of early-stage cancers using circulating tumor DNA. Sci Transl Med. 2017;9:2415.
57. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med. 2008;14:985–90.
58. Chen K, Zhang J, Guan T, Yang F, Lou F, Chen W, et al. Comparison of plasma to tissue DNA mutations in surgical patients with non-small cell lung cancer. J Thorac Cardiovasc Surg. 2017;154:1123-1131.e2.
59. Merker JD, Oxnard GR, Compton C, Diehn M, Hurley P, Lazar AJ, et al. Circulating tumor DNA analysis in patients with cancer: american society of clinical oncology and college of American pathologists joint review. J Clin Oncol. 2018;36:1631–41.
60. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med. 2014;371:2488–98.
61. Razavi P, Li BT, Brown DN, Jung B, Hubbell E, Shen R, et al. High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. Nat Med. 2019;25:1928–37.
62. Edge SB, Compton CC. The American joint committee on cancer: the 7th edition the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol. 2010;17:1471–4.

Publisher's Note

Supplemental Table 1. Statistical parameters of CTCs and CTECs of 104 participating subjects

Characteristics	CTCs (Units/6ml)									CTECs (Units/6ml)								
	Total CTCs			$_{L}$CTCs			sCTCs			Total CTECs			$_{L}$CTECs			sCTECs		
	M	**P25-P75**	***P***	**M**	**P25-P75**	***P***	M	**P25-P75**	***P***	**M**	**P25-P75**	***P***	**M**	**P25-P75**	***P***	**M**	**P25-P75**	***P***
Healthy	5.5	3-8.75	<0.001	3	1-6	0.027	2	0-4.5	<0.0001	6	2-10	<0.001	3	2-6.75	0.0001	0	0-0	<0.0001
BLDP	5	1-11		2	0-5		2	0-5		5	2-7		5	0-8		1	0-2	
NSCLC	16	7-22.5		5	1.5-11.5		10	3.5-17		18	11.5-28.5		13	5-24		4	2-8	
Pathological type			0.629			0.153			0.891			0.183			0.557			0.156
AC	16	8.25-21.75		4	2-10.75		10.5	3.25-17		17	9.5-27		14	5-23.75		4	1-1.75	
SC	14	5.5-37.75		7.5	2.25-26		8	4.75-12		20	13.25-36.5		12	4.5-30		5	3-8.75	
ASC	34	3-90		16	2-50		18	1-40		32	22-131		18	13-115		16	4-19	
Other	9	2-16		0.5	0-1		8.5	2-15		10.5	6-15		8	2-14		2.5	1-4	
TNM stage			0.715			0.747			0.564			0.388			0.525			0.826
I	16	6-23		4	1-12		13	2-17		17	8-30		11	4-26		4	2-8	
II	36	5-67		27.5	1-54		8.5	4-13		18	17-19		14	12-16		4	1-7	
III	14	7.5-17		5	1.5-8.5		7	3-9.5		17	7-23.5		12	1.5-19		4	0.5-8.5	
IV	16	14-34		5	4-16		12	8-18		25	16-73		18	9-67		6	2-8	

Measurement data were compared between two groups by Mann-Whitney U test, multiple groups by Kruskal-Wallis test. Abbreviations: NSCLC non-small cell lung cancer, BLDP benign lung disease patients, AC Adenocarcinoma, SC Squamous carcinoma, ASC adenosquamous carcinoma, CTCs circulating tumor cells, CTECs circulating tumor-derived endothelial cells, M median, P25-P75 inter-quartile range

Supplemental Table 2. Statistical parameters of aneuploid CTCs and CTECs of 104 participating subjects

Characteristics	Aneuploid CTCs (Units/6ml)									Aneuploid CTECs (Units/6ml)								
	Triploid			Tetraploid			Multiploid			Triploid			Tetraploid			Multiploid		
	M	**P25-P75**	***P***	**M**	**P25-P75**	***P***	**M**	**P25-P75**	***P***	**M**	**P25-P75**	***P***	**M**	**P25-P75**	***P***	**M**	**P25-P75**	***P***
Healthy	1	0-2	<0.001	0	0-1.75	<0.0001	3	1-6	0.058	0	0-0	<0.0001	0	0-0	<0.0001	3	0.25-6	<0.001
BLDP	2	0-5		1	0-3		1	0-4		1	0-3		0	0-3		4	1-8	
NSCLC	8	3.5-12		4	1.5-6		3	1-7.5		3	1-6		2	1-4		8	4-23	
Pathological type			0.887			0.657			0.428			0.069			0.083			0.127
AC	7.5	4-13.5		4	2-5.75		8	4-22		2.5	1-6		2	0-4		8.5	4.25-22	
SC	8	4-8.75		2	0.25-12		10.5	5.75-30		4	3-5.75		2.5	1.25-4		10.5	5.75-30	
ASC	11	1-40		6	1-9		12	8-106		8	8-8		8	6-11		12	8-106	
Other	6.0	1-11.0		2.5	1-4		1	1-1		1.5	0-3		1.5	1-2		3.5	1-6	
TNM stage			0.240			0.897			0.365			0.134			0.653			0.389
I	9	4-14		4	2-6		2	0-7		2	0-6		2	1-4		7	4-24	
II	4.5	3-6		12	1-23		19.5	1-38		2	1-3		2	0-4		7.5	3-12	
III	4	2-8		4	0-5.5		3	2-7.5		5	1-6		1	0-4		8	3-16	
IV	11	8-14		3	2-6		4	2-17		5	3-8		3	0-7		14	8-58	

The number of CTEC triploid in adenosquamous carcinoma was significantly higher than that in adenocarcinoma and squamous carcinoma (p=0.0161*, Mann-Whitney U test; p=0.0424*, Mann-Whitney U test). The number of CTEC tetraploid in adenosquamous carcinoma was significantly higher than that in adenocarcinoma and squamous carcinoma (p=0.0061**, Mann-Whitney U test; p=0.0061**, Mann-Whitney U test). The number of CTEC triploid in stage IV was significantly higher than that in stage I (p=0.0276*, Mann-Whitney U test). Measurement data were compared between multiple groups by Kruskal-Wallis test.

Abbreviations: NSCLC non-small cell lung cancer, BLDP benign lung disease patients, AC Adenocarcinoma, SC Squamous carcinoma, ASC adenosquamous carcinoma, CTCs circulating tumor cells, CTECs circulating tumor-derived endothelial cells, M median, P25-P75 inter-quartile range.

Supplemental Table3. The diagnostic sensitivity of CTCs and CTECs in different pathological types and stages of NSCLC

Characteristics	CTCs									CTECs								
	Total CTCs			$_{L}$CTCs			sCTCs			Total CTECs			$_{L}$CTECs			sCTECs		
	⩾11.5	<11.5	SEN%	⩾6.5	<6.5	SEN%	⩾6.5	<6.5	SEN%	⩾10.5	<10.5	SEN%	⩾8.5	<8.5	SEN%	⩾2.5	<2.5	SEN%
Pathological type																		
AC	27	9	75.0	12	24	33.3	23	13	63.9	27	9	75.0	23	13	63.9	22	14	61.1
SC	6	2	75.0	5	3	62.5	6	2	75.0	7	1	87.5	6	2	75.0	7	1	87.5
ASC	2	1	66.7	2	1	66.7	2	1	66.7	3	0	100.0	3	0	100.0	3	0	100.0
Other	1	1	50.0	0	2	0.0	1	1	50.0	1	1	50.0	1	1	50.0	1	1	50.0
TNM stage																		
I	19	8	70.4	10	17	37.0	16	11	59.3	19	8	70.4	17	10	63.0	18	9	66.7
II	1	1	50.0	1	1	50.0	1	1	50.0	2	0	100.0	2	0	100.0	2	0	100.0
III	7	2	77.8	4	5	44.4	6	3	66.7	7	2	77.8	5	4	55.6	5	4	55.6
IV	9	2	81.8	4	7	36.4	9	2	81.8	10	1	90.9	9	2	81.8	8	3	72.7

Abbreviations: NSCLC non-small cell lung cancer, AC Adenocarcinoma, SC Squamous carcinoma, ASC adenosquamous carcinoma, CTCs circulating tumor cells, CTECs circulating tumor-derived endothelial cells.

Supplemental Table 4. The diagnostic sensitivity of aneuploid CTCs and CTECs in different pathological types and stages of NSCLC

Characteristics	Aneuploid CTCs									Aneuploid CTECs								
	Triploid			Tetraploid			Multiploid			Triploid			Tetraploid			Multiploid		
	⩾6.5	<6.5	SEN%	⩾4.5	<4.5	SEN%	⩾6.5	<6.5	SEN%	⩾3.5	<3.5	SEN%	⩾2.5	<2.5	SEN%	⩾7.5	<7.5	SEN%
Pathological type																		
AC	19	17	52.8	12	14	33.3	9	27	33.3	16	20	44.4	15	11	41.7	19	14	52.8
SC	6	2	75.0	2	6	25.0	3	5	37.5	4	4	50.0	4	4	50.0	6	2	75.0
ASC	2	1	66.7	2	1	66.7	2	1	66.7	3	0	100.0	3	0	100.0	3	0	100.0
Other	1	1	50.0	1	1	50.0	0	2	0.0	0	2	0.0	0	2	0.0	0	1	0.0
TNM stage																		
I	16	9	59.3	10	17	37.0	8	19	29.6	11	16	40.7	11	16	40.7	13	14	48.1
II	0	2	0.0	1	1	50.0	1	1	50.0	0	2	0.0	1	1	50.0	1	1	50.0
III	3	6	33.3	2	7	22.2	2	7	22.2	6	3	66.7	3	6	33.3	5	4	55.6
IV	9	2	81.8	3	8	27.3	3	8	27.3	7	4	63.6	7	4	63.6	9	2	81.8

Abbreviations: NSCLC non-small cell lung cancer, AC Adenocarcinoma, SC Squamous carcinoma, ASC adenosquamous carcinoma, CTCs circulating tumor cells, CTECs circulating tumor-derived endothelial cells.

Supplemental table5. Target gene information of ctDNA panel

ABL1	*AKT1*	*AKT2*	*ALK*	*APC*	*AR*	*ARAF*
BRAF	*BRCA1*	*BRCA2*	*CCND1*	*CCND2*	*CCND3*	*CD274*
CDK4	*CDK6*	*CDKN2A*	*CSF1R*	*CTNNB1*	*DDR2*	*DPYD*
EGFR	*ERBB2*	*ESR1*	*EZH2*	*FBXW7*	*FGFR1*	*FGFR2*
FGFR3	*FLT1*	*FLT3*	*FLT4*	*GATA3*	*GNA11*	*GNAQ*
GNAS	*IDH1*	*IDH2*	*JAK2*	*JAK3*	*KDR*	*KEAP1*
KIT	*KRAS*	*MAP2K1*	*MAP2K2*	*MET*	*MLH1*	*MSH2*
MSH6	*MTOR*	*NF2*	*NFE2L2*	*NRAS*	*NTRK1*	*PDCD1LG2*
PDGFRA	*PDGFRB*	*PIK3CA*	*PIK3R1*	*PMS2*	*PTCH1*	*PTEN*
RAF1	*RB1*	*RET*	*RNF43*	*ROS1*	*SMAD4*	*SMO*
STK11	*TERT*	*TP53*	*TSC1*	*TSC2*	*UGT1A1*	*VHL*

Supplemental table 6. Target gene information of tDNA panel

ABCB1(MDR1)	*ABCC2*	*ABL1*	*ADH1B*	*AIP*	*AKT1*	*AKT2*
AKT3	*ALDH2*	*ALK*	*AMER1*	*APC*	*APEX1*	*AR*
ARAF	*ARID1A*	*ARID1B*	*ARID2*	*ARID5B*	*ASCL4*	*ASXL1*
ATF1	*ATIC*	*ATM*	*ATR*	*ATRX*	*AURKA*	*AURKB*
AXIN2	*AXL*	*B2M*	*BAD*	*BAI3*	*BAK1*	*BAP1*
BARD1	*BAX*	*BCL2*	*BCL2L1*	*BCL2L11*	*BCL3*	*BCR*
BIRC3	*BLM*	*BMPR1A*	*BRAF*	*BRCA1*	*BRCA2*	*BRD4*
BRIP1	*BTG2*	*BTK*	*BUB1B*	*CLLOR30*	*CASP8*	*CBL*
CBLB	*CCND1*	*CCNE1*	*CD274(PD-1)*	*CD74*	*CDA*	*CDC73*
CDH1	*CDK10*	*CDK12*	*CDK4*	*CDK6*	*CDK8*	*CDKN1A*
CDKN1B	*CDKN1C*	*CDKN2A*	*CDKN2B*	*CDKN2C*	*CEBPA*	*CEP57*
CHD4	*CHD8*	*CHEK1*	*CHEK2*	*CREBBP*	*CRKL*	*CSF1R*
CTCF	*CTLA4*	*CTNNB1*	*CUL3*	*CUX1*	*CXCL8*	*CXCR4*
CYLD	*CYP19A1*	*CYP2A13*	*CYP2A6*	*CYP2A7*	*CYP2B6*6*	*CYP2C19*2*
*CYP2C9*3*	*CYP2D6*	*CYP3A4*4*	*CYP3A5*	*CYSLTR2*	*DAXX*	*DDR2*
DENND1A	*DHFR*	*DICER1*	*DLL3*	*DNMT3A*	*DOT1L*	*DPYD*
DTL(CDT2)	*DUSP2*	*EGFR*	*EIF1AX*	*EP300*	*EPAS1*	*EPCAM*
EPHA2	*EPHA3*	*EPHA5*	*ERBB2*	*ERBB2IP*	*ERBB3*	*ERBB4*
ERCC1	*ERCC2*	*ERCC3*	*ERCC4*	*ERCC5*	*ESR1*	*ETV1*
ETV4	*ETV5*	*ETV6*	*EWSR1*	*EXT1*	*EXT2*	*EZH2*
EZR	*FANCA*	*FANCC*	*FANCD2*	*FANCE*	*FANCF*	*FANCG*
FANCI	*FANCL*	*FANCM*	*FAT1*	*FBXW7*	*FGF19*	*FGFR1*
FGFR2	*FGFR3*	*FGFR4*	*FH*	*FLCN*	*FLT1*	*FLT3*
FLT4	*FOXA1*	*FOXL2*	*FOXO3*	*FOXP1*	*FRG1*	*GATA1*
GATA2	*GATA3*	*GATA4*	*GATA6*	*GNA11*	*GNAQ*	*GNAS*
GR1N2A	*GRM3*	*GRM8*	*GSTM1*	*GSTM4*	*GSTP1*	*GSTT1*
HDAC1	*HDAC2*	*HDAC9*	*HGF*	*HLA-A*	*HMOX1*	*HNF1A*
HNF1B	*HRAS*	*HSPB1*	*IDH1*	*IDH2*	*IFNA6*	*IFNB1*
IFNE	*IFNG*	*IFNGR1*	*IFNGR2*	*IGF1R*	*IGF2*	*IKBKE*
IKZF1	*IL13*	*IL1A*	*IL7R*	*INPP4B*	*IRF1*	*IRF2*
ITGB6	*JAK1*	*JAK2*	*JAK3*	*JARID2*	*JUN*	*KDM5A*
KDR(VEFGR2)	*KEAP1*	*KIF1B*	*KIT*	*KITLG*	*KLLN*	*KMT2A*
KMT2B	*KMT2C*	*KMT2D*	*KRAS*	*LHCGR*	*LIG3*	*LIG4*
LIN28B	*LMO1*	*LRP1B*	*LYN*	*LZTR1*	*MALT1*	*MAP2K1*
MAP2K2	*MAP2K4*	*MAP3K1*	*MAP3K4*	*MAPK1*	*MAPK3*	*MAX*
MCL1	*MDM2*	*MDM4*	*MECOM*	*MED12*	*MEF2B*	*MEN1*
MET	*MGMT*	*MIF*	*MITF*	*MLH1*	*MLH3*	*MLLT1*
MLLT3	*MLL4*	*MMP1*	*MPL*	*MRE11A*	*MSH2*	*MSH6*
MTHFR	*MTOR*	*MTRR*	*MUC5B*	*MUTYH*	*MYC*	*MYCL*
MYCN	*MYD88*	*MYH9*	*NAT1*	*NBN*	*NCOR1*	*NEIL1*
NF1	*NF2*	*NFE2L2*	*NFKB1*	*NFKBIA*	*NKX2-1*	*NOS2*

NOS3	*NOTCH1*	*NOTCH2*	*NOTCH3*	*NPM1*	*NQO1*	*NRAS*
NRG1	*NSD1*	*NTHL1*	*NTRK1*	*NTRK2*	*NTRK3*	*NUTM1*
PAK2	*PAK3*	*PALB2*	*PALLD*	*PARK2*	*PARP1*	*PARP2*
PAX5	*PBRM1*	*PDCD1*	*PDCD1LG2*	*PDE11A*	*PDGFRA*	*PDGFRB*
PDK1	*PGR*	*PHOX2B*	*PIK3C3*	*PIK3CA*	*PIK3CD*	*PIK3R1*
PIK3R2	*PKHD1*	*PLAG1*	*PLCB4*	*PLK1*	*PMAIP1*	*PMS1*
PMS2	*PNKP(PNK)*	*POLD1*	*POLD3*	*POLE*	*POLH*	*POLT1*
PPARD	*PPP2R1A*	*PRDM1*	*PREX2*	*PRF1*	*PRKACA*	*PRKAR1A*
PRKCB	*PRKCI*	*PRKDC*	*PRSS1*	*PRSS3*	*PTCH1*	*PTEN*
PTK2	*PTPN11*	*PTPN13*	*QKI*	*RAC1*	*RAC3*	*RAD50*
RAD51	*RAD51B*	*RAD51C*	*RAD51D*	*RAD54L*	*RAD9A*	*RAF1*
RARA	*RARG*	*RASGEF1A*	*RB1*	*RCC1*	*RECQL4*	*RELA*
RELN	*RET*	*RHOA*	*RICTOR*	*RNF43*	*ROS1*	*RPTOR*
RRM1	*RUNX1*	*RUNX1T1*	*SBDS*	*SDC4*	*SDHA*	*SDHB*
SDHC	*SDHD*	*SEPT9*	*SERPINE1*	*SETBP1*	*SETD2*	*SF3B1*
SGK1	*SKP2*	*SLC34A2*	*SLC3A2*	*SMAD2*	*SMAD3*	*SMAD4*
SMAD7	*SMARCA4*	*SMARCB1*	*SMO*	*SOCS1*	*SOS1*	*SOX2*
SPOP	*SPRED1*	*SPRY4*	*SRC*	*SRSF2*	*SRY*	*STAG2*
STAT1	*STAT3*	*STK11*	*STMN1*	*SUFU*	*SUMO1*	*TACC3*
TAP1	*TAP2*	*TBK1*	*TEK*	*TEKT4*	*TERC*	*TERT*
TET2	*TFG*	*TGFB1*	*TGFBR2*	*THADA*	*TMEM127*	*TMEM167A*
TMPRSS2	*TNF*	*TNFAIP3*	*TNFRSF11A*	*TNFRSF14*	*TNFRSF19*	*TNFRSF1B*
TNFSF11	*TOP1*	*TOP2A*	*TP53*	*TP63*	*TPMT*	*TSC1*
TSC2	*TSHR*	*TTF1*	*TUBB3*	*TYMS*	*USAF1*	*UGT1A1*
UNG	*VAMP2*	*VEGFA*	*VHL*	*WAS*	*WISP3*	*WRN*
WT1	*XPA*	*XPC*	*XRCC1*	*XRCC2*	*XRCC3*	*XRCC4*
XRCC5	*YAP1*	*ZNF2*	*ZNF217*	*ZNF703*	-	-

Supplemental Table7. Mutational profile of ctDNA in enrolled 28 patients

Gene	All Subjects n (%)(N=28)	Pathological type	Variant CDSchange	Variant type	Variant Description	Alle Fraction(%)
TP53	9(32.1)	AC	exon4: c.374C>A: p.Thr125Lys	SNV	Missense	0.38
		AC	Intron7: c.919+1G>A: ……	SNV	Intronic	0.29
		SCC	exon6: c.646G>T: p.Val216Leu	SNV	Missense	0.12
		SCC	exon7: c.701A>G: p.Tyr234Cys	SNV	Missense	7.29
		SCC	exon6: c.592G>T: p.Glu198*	SNV	Nonsense	13.85
		SCC	exon6: c.583A>T: p.Ile195Phe	SNV	Missense	29.97
		SCC#	exon6: c.578A>G: p.His193Arg	SNV	Missense	7.23
		SCC#	exon7: c.711G>A: p.Met237Ile	SNV	Missense	3.88
		SCC	exon6: c.583A>T: p.Ile195Phe	SNV	Missense	0.52
		SCC	exon7: c.713G>A: p.Cys238Tyr	SNV	Missense	0.16
RB1	2(7.1)	Benign	exon19: c.1862G>A: p.Arg621His	SNV	Missense	48.9
		SCC	Intron20: c.2107-1G>A: ……	SNV	Intronic	13.3
ERBB2	2(7.1)	Benign	exon23: c.2810A>G: p.Lys937Arg	SNV	Missense	0.34
		SCC	exon23: c.2810A>G: p.Lys937Arg	SNV	Missense	0.37
IDH1	2(7.1)	Benign	exon4: c.394C>A: p.Arg132Ser	SNV	Missense	0.05
		AC	exon4: c.394C>G: p.Arg132Gly	SNV	Missense	2.16
PTEN	2(7.1)	Benign	exon8: c.962_963insA: p.Asn323fs	Delins	Frameshift	0.12
		SCC	Intron8: c.1026+1G>C: ……	SNV	Intronic	1.77
MET	2(7.1)	Benign	exon14: c.3013C>G: p.Arg1005Gly	SNV	Missense	0.09
		SCC	CNV SCORE:4.75	CNV	-	-
ALK	2(7.1)	AC	exon12: c.2073C>T: p.Ser691Ser	SNV	Synonymous	42.63
		SCC	exon28: c.4076A>G: p.Tyr1359Cys	SNV	Missense	6.85
FGFR2	2(7.1)	SCC	exon12: c.1650T>A: p.Asn549Lys	SNV	Missense	0.52
		SCC	exon9: c.1147T>C: p.Cys383Arg	SNV	Missense	0.08
RET	1(3.6)	AC	exon11: c.1891G>A: p.Asp631Asn	SNV	Missense	41.54
PTCH1	1(3.6)	AC	Intron9: c.1347+6G>A: ……	SNV	Intronic	45.65
DDR2	1(3.6)	AC	exon14: c.1530C>T: p.Val510Val	SNV	Synonymous	45.07
TSC2	1(3.6)	SCC	exon20: c.2153G>A: p.Arg718His	SNV	Missense	44.64
MSH6	1(3.6)	SCC	exon5: c.3197A>G: p.Tyr1066Cys	SNV	Missense	0.07
GNAS	1(3.6)	SCC	exon8: c.604C>T: p.Arg201Cys	SNV	Missense	0.92
NFE2L2	1(3.6)	SCC	exon2: c.92G>C: p.Gly31Ala	SNV	Missense	15.43
KEAP1	1(3.6)	SCC	exon4: c.1408C>T: p.Arg470Cys	SNV	Missense	15.15
EGFR	1(3.6)	SCC	CNV SCORE:3.79	CNV	-	-
MAP2K1	1(3.6)	SCC	exon6: c.644T>C: p.Leu215Pro	SNV	Missense	0.06
PIK3CA	1(3.6)	SCC	exon21: c.3140A>G: p.His1047Arg	SNV	Missense	0.11
FGFR3-TACC3	1(3.6)	SCC	Fusion	-	-	-

The genes name in bold means in the list of 75 overlapping genes between ctDNA and tDNA panel,

Two different mutations in the same patient

Supplemental Table8. Mutational profiles of tumor DNA in filtered 9 patients

Gene	All Subjects n (%) (N=9)	Pathological type	Variant CDSchange	Variant type	Variant Description	Alle Fraction(%)
EGFR	5(55.6)	AC	exon21:c.2573T>G: p.L858R	SNV	Missense	55.2
		AC	exon21:c.2573T>G: p.L858R	SNV	Missense	15.3
		AC	exon21:c.2573T>G: p.L858R	SNV	Missense	23.1
		AC	exon21:c.2573T>G: p.L858R	SNV	Missense	21.2
		SCC	exon19:c.2237_2255delinsT:p.E746_S752delinsV	Delins	nonFrameshift	39.5
MCL1	4(44.4)	AC	2.6 copies	CNV	-	-
		AC	2.6 copies	CNV	-	-
		AC	6.4 copies	CNV	-	-
		AC	2.4 copies	CNV	-	-
CDK4	3(33.3)	AC	7.6 copies	CNV	-	-
		AC	2.1 copies	CNV	-	-
		AC	2.3 copies	CNV	-	-
VEGFA	2(22.2)	AC	7.4 copies	CNV	-	-
		AC	2.2 copies	CNV	-	-
BCL3	2(22.2)	AC	9.8 copies	CNV	-	-
		AC	2.3 copies	CNV	-	-
FGFR1	2(22.2)	AC	exon13:c.1727G>A: p.R576Q	SNV	Missense	63.0
		SCC	exon5:c.565C>T:p.R189C	SNV	Missense	9.9
MYC	2(22.2)	AC	2.2 copies	CNV	-	-
		SCC	2.1 copies	CNV	-	-
PDGFRB	2(22.2)	AC	exon7: c.1084G>A: p.A362T	SNV	Missense	1.3
		SCC	exon15: c.2129C>T:p.P710L	SNV	Missense	8.8
CYLD	2(22.2)	SCC	exon4:c.173C>G:p.S58*	SNV	Nonsense	34.1
		SCC	exon10:c.1145A>T:p.E382V	SNV	Missense	2.6
TP53	2(22.2)	SCC	exon6:c.646G>T:p.V216L	SNV	Missense	36.5
		SCC	exon8:c.833C>T:p.P278L	SNV	Missense	39.0
FGFR2	1(11.1)	SCC	exon3:c.290C>T:p.A97V	SNV	Missense	9.0
DUSP2	1(11.1)	Benign	exon4: c.914T>A: p.L305Q	SNV	Missense	46.4
CDA	1(11.1)	Benign	exon2: c.160A>T: p.N54Y	SNV	Missense	47.2
RB1	1(11.1)	Benign	exon19: c.1862G>A:p.R621H	SNV	Missense	49.8
RET	1(11.1)	AC	exon11: c.1891G>A:p.D631N	SNV	Missense	41.9
ARID1A	1(11.1)	AC	exon18:c.4798_4799delTCinsAA: p.S1600N	Delins	Missense	13.1
CUL3	1(11.1)	AC	exon9: c.1288C>T: p.R430C	SNV	Missense	35.4
MLH1	1(11.1)	AC	exon12:c.1404C>A: p.N468K	SNV	Missense	62.2
STMN1	1(11.1)	AC	2.7 copies	CNV	-	-
MDM4	1(11.1)	AC	2.5 copies	CNV	-	-
CBL	1(11.1)	AC	exon9: c.1384C>T: p.R462*	SNV	Nonsense	2.1
POLD1	1(11.1)	AC	Single copy	CNV	-	-
GATA6	1(11.1)	AC	exon2: c.145G>A: p.G49S	SNV	Missense	35.6

NSD1	1(11.1)	AC	exon23: c.7493G>A: p.G2498E	SNV	Missense	22.1
RAD50	1(11.1)	AC	exon21: c.3278G>T: p.R1093L	SNV	Missense	21.2
MDM2	1(11.1)	AC	2.8 copies	CNV	-	-
TOP1	1(11.1)	AC	exon15:c.1517C>G: p.S506*	SNV	Nonsense	2.4
MLH1	1(11.1)	AC	exon12:c.1404C>A: p.N468K	SNV	Missense	36.2
BMPR1A	1(11.1)	SCC	Single copy	CNV	-	-
PALLD	1(11.1)	SCC	exon10: c.1847_1859delGCCGTG GAGTAAA: p.S616Mfs*3	Delins	Frameshift	32.0
PRKAR1A	1(11.1)	SCC	exon2:c.46C>T:p.R16*	SNV	Nonsense	15.9
PTEN	1(11.1)	SCC	exon9:c.1030A>T:p.K344*	SNV	Nonsense	34.2
WISP3	1(11.1)	SCC	exon2:c.220C>T:p.Q74*	SNV	Nonsense	3.3
ATR	1(11.1)	SCC	exon14:c.2963A>G:p.N988S	SNV	Missense	20.8
CBLB	1(11.1)	SCC	exon5:c.607G>A:p.K7E203K	SNV	Missense	2.1
DOT1L	1(11.1)	SCC	exon20:c.2327C>G:p.P776R	SNV	Missense	31.4
FANCC	1(11.1)	SCC	exon5:c.406C>G:p.Q136E	SNV	Missense	2
GATA2	1(11.1)	SCC	exon2:c.206G>T:p.R69L	SNV	Missense	21.3
IFNGR1	1(11.1)	SCC	exon7:c.1187C>G:p.S396W	SNV	Missense	4.0
MED12	1(11.1)	SCC	exon15:c.2177A>C:p.Y726S	SNV	Missense	2.0
NPM1	1(11.1)	SCC	exon10:c.844G>A:p.E282K	SNV	Missense	34.0
ROS1	1(11.1)	SCC	exon4:c.265_266delGGinsTT:p.G89F	Delins	Missense	40.3
WRN	1(11.1)	SCC	exon26:c.3194C>T:p.A1065V	SNV	Missense	3.1
HARS	1(11.1)	SCC	exon2:c.35G>A:p.G12D	SNV	Missense	17.3
LRP1B	1(11.1)	SCC	exon78:c.11952C>G:p.Y3984*	SNV	Nonsense	27.2
ARID1B	1(11.1)	SCC	exon12:c.3185A>G:p.E1062G	SNV	Missense	31.5
AXIN2	1(11.1)	SCC	exon2:c.325G>A:p.D109N	SNV	Missense	10.2
BRIP1	1(11.1)	SCC	exon17:c.2387T>C:p.L796P	SNV	Missense	37.8
EPHA3	1(11.1)	SCC	exon14:c.2362A>T:p.I788F	SNV	Missense	37.7
FANCL	1(11.1)	SCC	exon14:c.1124A>C:p.H375P	SNV	Missense	2.7
FLT3	1(11.1)	SCC	exon12:c.1435A>G:p.T479A	SNV	Missense	35.6
FLT4	1(11.1)	SCC	exon15:c.2228C>T:p.A743V	SNV	Missense	2.0
TSC2	1(11.1)	SCC	exon20: c.2153G>A:p.R718H	SNV	Missense	47.8
LYN	1(11.1)	SCC	exon13:c.1372C>T:p.Q458*	SNV	Nonsense	27.7
MYCN	1(11.1)	SCC	exon3:c.1172G>T:p.R391L	SNV	Missense	22.4
PAK3	1(11.1)	SCC	exon14:c.1274C>A:p.T425N	SNV	Missense	10.7
EP300	1(11.1)	SCC	exon24:c.3939_3940delGA:p.N1314Sfs*3	Delins	Frameshift	11.1
GNAS	1(11.1)	SCC	2.1 copies	CNV	-	-
QKI	1(11.1)	SCC	exon3:c.401delA:p.K134Rfs*14	Delins	Frameshift	8.4
SMARCA4	1(11.1)	SCC	exon19:c.2729C>T:p.T910M	SNV	Missense	9.5
SPRY4	1(11.1)	SCC#	exon3:c.579_580delTG:p.A194Ifs*42	Delins	Frameshift	10.3
		SCC#	exon3:c.113C>A:p.S38*	SNV	Nonsense	1.0
ABL1	1(11.1)	SCC#	exon10:c.1651C>T:p.P551S	SNV	Missense	4.2
		SCC#	exon11:c.2770G>A:p.G924R	SNV	Missense	1.0
AKT2	1(11.1)	SCC	exon10:c.839A>C:p.N280T	SNV	Missense	2.6
CSF1R	1(11.1)	SCC	exon4:c.363G>T:p.Q121H	SNV	Missense	9.0

ETV4	1(11.1)	SCC	exon7:c.455A>G:p.Q152R	SNV	Missense	8.0
GRM3	1(11.1)	SCC	exon4:c.2166G>T:p.K722N	SNV	Missense	1.5
LZTR1	1(11.1)	SCC	exon18:c.2173C>T:p.R725C	SNV	Missense	2.8
POLE	1(11.1)	SCC	exon8:c.743A>G:p.Y248C	SNV	Missense	8.3
SEPT9	1(11.1)	SCC	exon2:c.547G>A:p.A183T	SNV	Missense	1.2
THADA	1(11.1)	SCC	exon33:c.4868C>T:p.T1623M	SNV	Missense	1.4
XRCC2	1(11.1)	SCC	exon3:c.404G>A:p.S135N	SNV	Missense	33.1
YAP1	1(11.1)	SCC	exon6:c.1058A>G:p.Y353C	SNV	Missense	1.6
CYP2A6	1(11.1)	SCC	exon1:c.100C>T:p.P34S	SNV	Missense	2.1
PNKP	1(11.1)	SCC	exon2:c.53delG:p.G18Efs*21	Delins	Frameshift	6.4
LIG3	1(11.1)	SCC	exon3:c.643G>A:p.A215T	SNV	Missense	7.5
RELA	1(11.1)	SCC	exon10:c.989G>A:p.R330H	SNV	Missense	1.4

The genes name in bold means in the list of 75 overlapping genes between ctDNA and tDNA panel,

Two different mutations in the same patient

Supplemental Table 9. Target gene information of HRR panel

AR	*ATM*	*ATR*	*BARD1*	*BRCA1*	*BRCA2*	*BRIP1*
CDH1	*CDK12*	*CHEK1*	*CHEK2*	*ESR1*	*FANCA*	*FANCL*
HDAC2	*HOXB13*	*MRE11A*	*NBN*	*PALB2*	*PPP2R2A*	*PTEN*
RAD51B	*RAD51C*	*RAD51D*	*RAD54L*	*STK11*	*TP53*	
BRAF	*ERBB2*	*KRAS*	*NRAS*	*PIK3CA*		

The genes from *AR* to *TP53* are related to the homologous recombination repair (HRR) pathway, covering the coding region and the exon-intron junction region. The coverage range of extra five important tumor-related driver genes (*BRAF ERBB2 KRAS NRAS* and *PIK3CA)* is as follows: *BRAF*, 11/12/15/18 exon hotspot regions; *ERBB2*, 1/3/8/9/11/16~21/23/25/27 exon hotspot regions; *KRAS&NRAS*, 2/3/4 exon hotspot regions; *PIK3CA*, 2/5/6/8/10/14/21 exon hotspot regions.

Supplemental Table 10. Target gene information of 61 GENE panel

APC	*ATM*	*BARD1*	*BMPR1A*	*BRAF*	*BRCA1*	*BRCA2*	*BRIP1*	*CDH1*
CDK4	*CDKN2A*	*CHEK2*	*ELAC2*	*EPCAM*	*FANCC*	*FH*	*FLCN*	*GNAS*
GREM1	*HOXB13*	*HRAS*	*KIT*	*MAX*	*MEN1*	*MET*	*MLH1*	*MRE11*
MSH2	*MSH6*	*MUTYH*	*NBN*	*NF1*	*NTRK1*	*PALB2*	*PALLD*	*PDGFRA*
PMS2	*PRKAR1A*	*PTCH1*	*PTEN*	*RAD50*	*RAD51*	*RAD51C*	*RAD51D*	*RB1*
RET	*SDHA*	*SDHAF2*	*SDHB*	*SDHC*	*SDHD*	*SMAD4*	*SMARCA4*	*SMARCB1*
STK11	*TMEM127*	*TP53*	*TSC1*	*TSC2*	*VHL*	*WRN*		

The entire exons of 61 cancer-related genes in peripheral blood, as well as single nucleotide variants (SNVs) and short fragment insertion/deletion variants (INDELs) within the classic splice sites, were covered.

Supplemental Table 11. Mutational profiles of WBCs DNA in benign 4 patients

Gene	All Subjects n (%)(N=4)	Pathological type	Variant CDSchange	Variant type	Variant Description	Alle Fraction(%)	ACMG Classification
RB1	1(25%)	Benign	exon19: c.1862G>A: p.Arg621His	SNV	Missense	47.53	Uncertain
CHEK2	1(25%)	Benign	intron10:c.1095+12G>A:p.?	Intronic	-	44.26	Uncertain
CDK12	1(25%)	Benign	exon14:c.4232C>T:p.(Ala1411Val)	SNV	Missense	51.22	Uncertain

The genes name in bold means in the list of 75 overlapping genes between ctDNA and tDNA panel

Journal of Cancer Research and Clinical Oncology (2022) 148:2681–2692
https://doi.org/10.1007/s00432-021-03852-8

ORIGINAL ARTICLE – CLINICAL ONCOLOGY

Detection of circulating rare cells benefitted the diagnosis of malignant solitary pulmonary nodules

Jianzhu Xie[1] · Zheng Ruan[2] · Jian Zheng[2] · Yanping Gong[1] · Yulan Wang[1] · Binjie Hu[1] · Jin Cheng[1] · Qian Huang[1]

Received: 21 August 2021 / Accepted: 5 November 2021 / Published online: 17 November 2021

Abstract

Introduction Solitary pulmonary nodules (SPNs) are challenging in differentiating between benignancy and malignancy. Therefore, more effective non-invasive biomarkers are urgently needed. The purpose of this investigation was to examine whether circulating rare cells (CRCs) could facilitate the differentiation between benign and malignant SPNs as well as its sensitivity and specificity.

Methods 164 patients diagnosed with SPNs, 24 healthy volunteers, and 25 patients diagnosed with advanced-stage lung cancer were included. CT/PET-CT images, serum tumor markers, and biopsy results were collected. The CRCs were examined using subtraction enrichment and immunostaining-fluorescence in situ hybridization (SE-iFISH) and their relationship with malignant or benign SPNs was analyzed.

Results The total CRC numbers from patients with malignant SPNs diagnosed by biopsy were significantly greater compared to those with benign SPNs ($P < 0.0001$), but not significantly different from patients with advanced lung cancer ($P > 0.05$). The total CRCs, with a cut-off value of 21.5 units, showed 67.6% sensitivity and 73.3% specificity [area under curve (AUC) 95% CI, 0.778 (0.666–0.889)] in discriminating benign and malignant SPNs and the triploid CRCs exhibited a high positive likelihood ratio of 8.4, which suggested that CRCs appeared to have a distinct advantage in discriminating benign and malignant SPNs compared to CT/PET-CT images and serum tumor markers and could be a potential screening indicator for lung cancer in the high-risk population.

Conclusions SE-iFISH could effectively detect CRCs including circulating tumor cells (CTCs) and circulating tumor-derived endothelial cells (CTECs) and the detection of CRCs could benefit the differentiation of patients with benign and malignant SPNs.

Keywords Solitary pulmonary nodule · Circulating rare cells · Circulating tumor cells · Circulating tumor-derived endothelial cells · SE-iFISH

Abbreviations

SPN	Solitary pulmonary nodule
CRCs	Circulating rare cells
CTCs	Circulating tumor cells
CTECs	Circulating tumor-derived endothelial cells
SE-iFISH	Subtraction enrichment and immunostaining-fluorescence in situ hybridization

Introduction

The inability to diagnose lung cancer early has continued to be a major challenge in the clinic, based on the early diagnosis rate of only 0.1–0.17% (Kasmann et al. 2017). Solitary pulmonary nodules (SPNs) are the most common pulmonary lesion detected by CT scanning and could be a sign of early-stage lung cancer. The prevalence of non-calcified SPNs has been reported to range from 8.5 to 21.2% (Gomez-Saez et al. 2014; Hammerschlag et al. 2015; Ju et al. 2013), as well as 26.3% in the high-risk population (National

✉ Jin Cheng
ajinch@163.com

✉ Qian Huang
huangqian_sjtu@163.com

[1] Molecular Diagnostics Laboratory of Cancer Center, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2] Department of Thoracic Surgery, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Lung Screening Trial Research et al. 2011). The diameter of pulmonary nodules is closely related to the degree of malignancy, McWilliams et al. (2013) showed that the malignancy rate of nodules with a diameter ≤ 6 mm was 0–1%, a diameter of 6–10 mm was 1–4%, and a diameter of 10–20 mm was 4–15%. In addition, the nature of the nodules is also important for the determination of benign versus malignant nodules. Studies have shown that the malignancy rate of pure ground-glass nodules was 59–73%, which was higher than that of purely solid nodules (7–9%) (Henschke et al. 2002). To date, low-dose CT (LDCT) imaging is still the first choice for SPN screening, and surgical biopsy is the only means of confirming benign or malignant nodules but is not recommended for solitary pulmonary nodules smaller than 8 mm in diameter (Detterbeck et al. 2013; Wood et al. 2018). Other ancillary diagnostic methods, including serological detection of tumor biomarkers related to lung cancer, such as squamous cell carcinoma antigen (SCC), cancer embryonic antigen (CEA), and cytokeratin-19 fragment (CYFRA 21-1), are often less specific because their expression levels may be elevated due to other diseases or inflammatory responses. As to invasive biopsies, poor compliance is the major drawback.

Recently, the use of liquid biopsy technology for the early diagnosis of cancer, which focuses on patients' body and blood fluid samples, has received increased attention. The fluid samples are analyzed for circulating rare cells (CRCs), cell-free circulating tumor DNA(ctDNA) that are shed from tumor cells, platelets, and exosomes (Ramalingam and Jeffrey 2018; Revelo et al. 2019). Thus, liquid biopsy sampling and analysis are non-invasive, less complicated to attain samples than tissue biopsies, and is more conducive to monitoring disease progression in real-time in the clinic. Several investigations have described a significant increase in concentrations of tumor-associated liquid biopsy markers in peripheral blood samples obtained from patients with prostate (Zhang and Armstrong 2016) and colon carcinomas (De Luca et al. 2000), as well as metastatic breast cancer (Bidard et al. 2016).

The circulating tumor cells (CTCs), defined as tumor cells that are shed from either solid primary or metastatic tumor tissues into the blood circulatory system (Lin 2018), are the most common subset of CRCs and have been proposed to serve as new markers for breast cancer by the American Society of Clinical Oncology (ASCO) since 2007 (Harris et al. 2007). Numerous clinical studies have confirmed that CTCs can serve as a robust and accurate diagnostic tool to monitor cancer progression and metastasis efficiently, and determine the prognosis for lung cancer (Guo Qiaomei 2016; Li et al. 2020; Wang et al. 2020; Zhou et al. 2017). Furthermore, a new subset of circulating rare cells that are CD31 positive and named circulating tumor-derived endothelial cells (CTECs) has been found in the circulating blood of lung cancer patients, as well as the significance with respect to the prognosis of non-small cell lung cancer (NSCLC) immunotherapy (Lin et al. 2017; Zhang et al. 2020).

A wide variety of CTC detection techniques (Grover et al. 2014; Mikolajczyk et al. 2011; Tartarone et al. 2017; Yu et al. 2011) have been reported. The most commonly used techniques are based on selectively capturing cells with antigens derived from epithelial cells or staining for various epithelium-derived proteins, such as cytokeratin (CK) as well as epithelial cell adhesion molecule (EpCAM). However, the expression of EpCAM on tumor cells is dynamic and highly heterogeneous (Went et al. 2004). In addition, the expression of CK is down-regulated during the transition from the epithelial state to mesenchyme (EMT) (Kalluri 2009). Therefore, techniques that are based on the expression of CK or EpCAM are likely to fail capture in some above-mentioned instances. In addition, these techniques also miss the detection of CTECs without anti-CD31 antibodies. More comprehensive, stable, and easier operated detection techniques are urgently needed in clinical practice. Here we reported that a novel approach, which integrated antigen-independent subtraction enrichment with the staining of multiple fluorescent-labelled antibodies and in situ hybridization (SE-iFISH), was used to detect peripheral blood CTCs and CTECs in patients and explore its potential role in identifying malignant or benign SPNs.

Materials and methods

Subjects and specimens

This was a retrospective investigation that included 164 patients who had SPNs. The investigation was carried out at the Department of Thoracic Surgery in Shanghai General Hospital, from July 2016 to April 2019. The patients in this study were diagnosed with SPNs by two radiologists using LDCT imaging. Patients with nodule calcification and multiple pulmonary nodules were excluded. 109 patients received surgical resection as suggested by published guidelines (Detterbeck et al. 2013; Wood et al. 2018), Based on the postoperative pathologic diagnosis. 60 patients had malignant SPNs and 49 patients had benign SPNs. 55 patients who were recommended for follow-up without surgical intervention were included in the non-operated group. The controls included 24 healthy volunteers and 25 patients with advanced lung cancer. All blood samples were collected before operation or chemotherapy.

Peripheral venous blood samples were collected using ACD vacuum tubes that contained anticoagulants. The epithelial cell contamination by skin that could be occurred during insertion of the needle was avoided by discarding the initial 2 ml of the sampled blood, then 6 ml of blood was collected in a new ACD vacuum tube. The blood was

immediately mixed by inverting the tube up and down eight times. The samples were stored away from light at room temperature (RT), and detection was performed within 24 h after sample collection. This study was performed in accordance with the Declaration of Helsinki and approved by the Ethical Review Board of Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, China (2016KY130). Informed consent was provided to all individuals in this study before blood collection.

Subtraction enrichment (SE) of CRCs

CRCs in 6 ml peripheral blood samples from all participants were enriched based on the updated manufacturer's protocols (Cytelligen, San Diego, CA, USA). At first, the plasma was removed by centrifuged (200 g) for 15 min at RT. The 1 × resuspension buffer was added to packed blood cells to achieve a total volume of 6 ml. Then this mixture was loaded onto the surface of a cell separation matrix that was non-hematopoietic in a 50 ml tube. Next, the samples were centrifuged (350 g) for 6 min to layer the different components in the mixture. The layer that contained the tumor cells as well as white blood cells (WBCs), which was located on top of the red blood cells (RBCs), was removed and placed into a new 50 ml tube with 300 µl of immuno-magnetic beads coated with antibodies against the leukocyte surface antigen (anti-CD45). The cells and immuno-magnetic beads were incubated for 30 min at RT. The WBCs that had been captured by immuno-magnetic beads were removed using a magnetic separator (Cytelligen, San Diego, CA, USA). The solution, which no longer contained the immuno-magnetic beads, was transferred into another 50 ml tube, and 1 × resuspension buffer was added to obtain a total volume of 50 ml. Next, the samples were spun (500 g) for 5 min at RT. Then the supernatant was aspirated from the tube to achieve a final volume of 100 µl with the cells remaining in the bottom of the tube. The sedimented cells were resuspended gently. Finally, the cells underwent multiple immunofluorescence staining in suspension and in situ hybridization with a probe for chromosome 8 centromere (CEP8) on the slide.

iFISH and CRCs imaging

Five-channel double-marker-iFISH was performed based on the updated manufacturer's instructions (Cytelligen, San Diego, CA, USA). Three types of tumor biomarkers were selected to label a range of CTCs, including two phenotypic epithelial biomarkers (EpCAM or CK18) as well as programmed death-ligand 1 (PD-L1), anti-CD45 antibody was used to mark residual WBCs, and anti-CD31 antibody was used to identify circulating tumor endothelium-derived cells or phenotypic endothelial cells (CTECs). All the antibodies were labeled with specific post-fluorescence tags, including Alexa Fluor 488-CD31, -CK18, -EpCAM, or -PD-L1 (green), Alexa Fluor 594-CD45 (red), as well as Cy5-CD31, -CK18, -EpCAM (gold) and worked at a dilution of 1:200 for 20 min in the dark. After rinsing, the cell monolayer that was dried was mixed with a specialized fixative that was then coated onto the formatted CRC slides (Cytelligen, San Diego, CA, USA). Subsequently, in situ hybridization was carried out on slides for four hours using the Vysis probe for chromosome 8 centromere (CEP8) Spectrum (Orange) on an S500 ThermoBrite Slide Hybridization System (Abbott, Abbott Park, IL, USA). Then a mounting medium that contained DAPI (Blue) (Vector Laboratories, Burlington, CA, USA) was applied to the slide. The stained samples were manual observed, identified, and photographed using fluorescence microscopy or were automatically scanned and imaged by CRC Metafer-Image Scanning System (Carl Zeiss, Oberkochen, Germany; MetaSystems, Altlussheim, Germany; and Cytelligen, San Diego, CA, USA), which could recognize five to six-color and have high throughput scanning and images.

Identification of CTCs and CTECs

Several criteria were used to define interested cells as positive. The essential morphological criteria included cells with intact nucleus. The general size of WBCs was used as a measure to define small cells (≤ the size of WBCs) or large cells (> the size of WBCs). The cells with interest were determined to be CTCs based on the observed expression of at least one of the tumor markers, either EpCAM, CK18, or PD-L1. The aneuploid karyotype and the increased size of the nucleus or its specific morphology were also taken into account during judgment. CTECs were determined by the observed expression of CD31 combined with aneuploid karyotype. It's very important to mention that cells that showed CD45 positive were excluded and would not be counted as CTCs or CTECs. The specific identification criteria are as follows: (1) CTC: DAPI+/CD45−/tumor biomarker+/CD31− with chromosome 8 diploid or aneuploid; DAPI+/CD45−/tumor biomarker−/CD31− with chromosome 8 aneuploid. (2) CTEC: DAPI+/CD45−/tumor biomarker±/CD31+ with chromosome 8 aneuploid. (3) CTM (circulating tumor microemboli): clusters of cells that contained two or more CTCs. In this study, the CRCs were identified initially through the use of manual identification with a fluorescence microscope for 87 subjects and the Metafer-image scanning assay for 126 subjects. It is worth to mention that all the CRCs images obtained from the Metafer-Image Scanning System were further reviewed and confirmed

by another two experienced and certified laboratory technicians.

Detection of tumor markers in serum

Three serum biological markers which were usually detected and recognized to be helpful for the diagnosis of lung cancer were included in this study, including squamous cell carcinoma antigen (SCC), cancer embryonic antigen (CEA), and cytokeratin-19 fragment (CYFRA 21-1). A chemiluminescent microparticle immunoassay and an Architect i2000 analyzer (Abbott) were used to analyze SCC. CYFRA21-1 and CEA were detected using electrochemical luminescence immunoassay (Roche Cobas® C-601). The procedures were precisely carried out following the kit instructions. The recommended reference indices that were used for the kits were as follows: SCC < 2.5 ng/ml, CEA < 5 ng/ml, and CYFRA21-1 < 3.3 ng/ml. A positive result was indicated when any one of these three indicators was observed to be above the normal value.

Statistical analysis

SPSS software (version 25.0) was utilized to process all data for analysis. Measurement data were compared between two groups using the Mann–Whitney *U* test or the two-tailed Student's *t*-test. The Kruskal–Wallis test was used to assess multiple groups. The inspection level was set at $\alpha = 0.05$. Correlation between the CRC numbers and the maximum diameter of the tumors was analyzed using the Pearson correlation coefficient. Analysis with ROC was used to identify the appropriate cut-off values, as well as evaluate the differential diagnostic efficiency of CRCs and other diagnostic indicators. *P*-values that were two-sided were used. A *P*-value less than 0.05 was determined to indicate statistical significance. Presentation of the statistical results was carried out using Prism 7.0 (GraphPad Software Inc., San Diego, CA, USA) as well as R, version 3.5.0.

Results

General information of research participants

A total of 213 research participants were enrolled from July 2016 to April 2019. The demographic characteristics of the enrolled participants and their groups are listed in Table 1. About 10 patients underwent twice or more detections but this study only accounted for the results from the first detection when patients had not received either surgery or other anti-tumor treatments. It is worth mentioning that most patients in the malignant SPN group exhibited adenocarcinoma pathologic characteristics, which accounted for 88.3% of the cases. In the lung adenocarcinoma subgroup, micro-invasive adenocarcinoma was the most common and accounted for 62.3% of the cases.

Table 1 Clinical characteristics of enrolled 213 research participants

Characteristics	No. (%)
Malignant SPNs (*N* = 60)	
Age years, median (range)	63.5 (22–80)
Sex (male/female)	22/38
Smoking (Y/N)	5/55
MTD mm, mean (range)	12.5 (3–30)
Pathological type	
AC	53 (88.3%)
ACIS	3 (5.7%)
MIA	33 (62.3%)
ADI	17 (32.1%)
SC	3 (5%)
ASC	2 (3%)
Other	2 (3%)
Benign SPNs (*N* = 49)	
Age years, median (range)	59 (21–73)
Sex (male/female)	23/26
Smoking (Y/N)	2/47
MTD mm, mean (range)	11.4 (3–30)
Non-operated SPNs (*N* = 55)	
Age years, median (range)	57 (33–79)
Sex (male/female)	30/25
Smoking (Y/N)	4/51
MTD mm, mean (range)	7.8 (2–30)
Advanced lung cancer (*N* = 25)	
Age years, median (range)	67 (46–80)
Sex (male/female)	21/4
Smoking (Y/N)	2/23
Pathological type	
AC	16 (64%)
SC	5 (4%)
Other	4 (16%)
TNM stage	
III	14 (56%)
IV	11 (44%)
Healthy volunteers (*N* = 24)	
Age years, median (range)	29 (23–60)
Sex (male/female)	8/16
Smoking (Y/N)	1/23

SPNs, solitary pulmonary nodules; MTD, maximum tumor diameter; AC, adenocarcinoma; SC, squamous carcinoma; ASC, adenosquamous carcinoma; ACIS, adenocarcinoma in situ; MIA, microinvasive Adenocarcinoma); ADI, adenocarcinoma infiltrating

Characterization of CRCs using SE-iFISH

Double-marker-iFISH was used to identify karyotypic and phenotypic characterization of CRCs at the same time since the Zeiss fluorescence microscope has 5–7 channels and can collect 5–7 signals from different antibodies staining and in situ hybridization. The criteria used to identify CTCs and CTECs have been described in the Materials and Methods section. A number of typical CTC cells (CD31−/CD45− aneuploidy cells or tumor marker+ cells) and CTEC cells (CD31+/CD45− aneuploidy cells) identified in this study were shown in Fig. 1.

The comparison of CRCs between manual observation and the Metafer-image scanning assay

In this study, 87 subjects from a total of 213 subjects, including 38 subjects from the non-operated SPN group, 19 subjects from the pathologic benign SPN group, 26 subjects from the pathologic malignant SPN group, and four subjects from the advanced lung cancer group were evaluated by manual observation. The remaining 126 subjects were evaluated by the Metafer-Image Scanning System followed by manual review and confirmation.

The CRC numbers in the group with malignant pulmonary nodules were significantly higher than that in the group with benign pulmonary nodules ($P < 0.05$, Fig. 2A-a; $P < 0.0001$, Fig. 2B-a). Furthermore, regardless of CRC size,

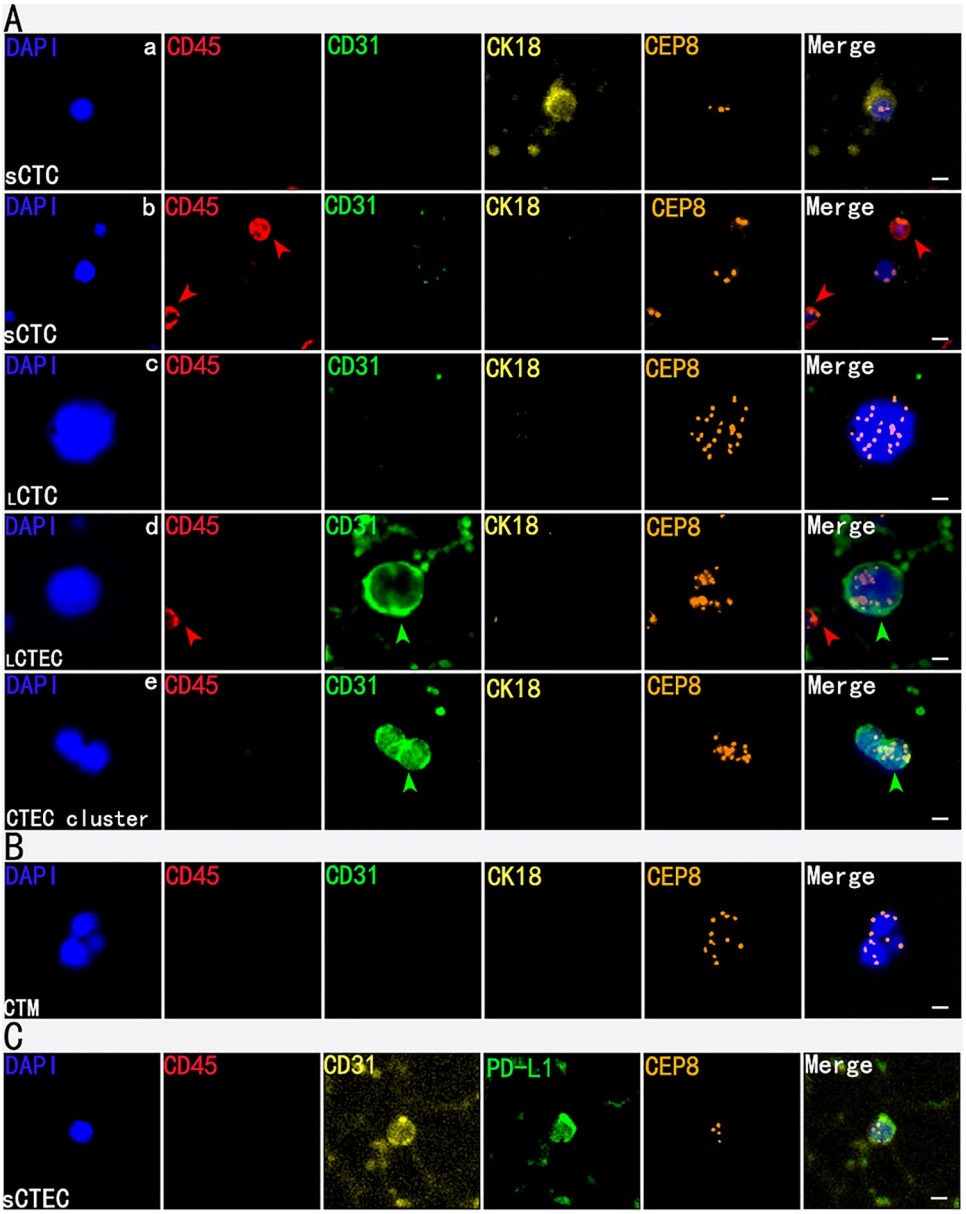


Fig. 1 Multi-fluorescence of CRCs: In situ single-tumor-biomarker SE-iFISH (five channel): nucleus (blue), CD45(red), CD31 (green/yellow), CK18 (yellow), PD-L1 (green) and CEP8 (orange). **A** CTCs and CTECs in patients with malignant solitary pulmonary nodules. **A-a**. DAPI+/CD45−/CD31−/CK18+/CEP8 small triploid CTC. **A-b** DAPI+/CD45−/CD31−/CK18−/CEP8 small triploid CTC with adjacent CD45+ WBCs (red arrow). **A-c** DAPI+/CD45−/CD31−/CK18−/CEP8 large multiploid CTC. **A-d** DAPI+/CD45−/CD31+/CK18−/CEP8 large multiploid CTEC (green arrow) with adjacent CD45+ WBC (red arrow). **A-e** DAPI+/CD45−/CD31+/CK18−/CEP8 multiploid CTECs cluster (green arrow). **B** DAPI+/CD45−/CD31−/CK18− CTM in patients with malignant solitary pulmonary nodules. **C** DAPI+/CD45−/CD31+/PD-L1+/CEP8 small multiploid CTEC in patients with advanced lung cancer. Bars: 5 μm

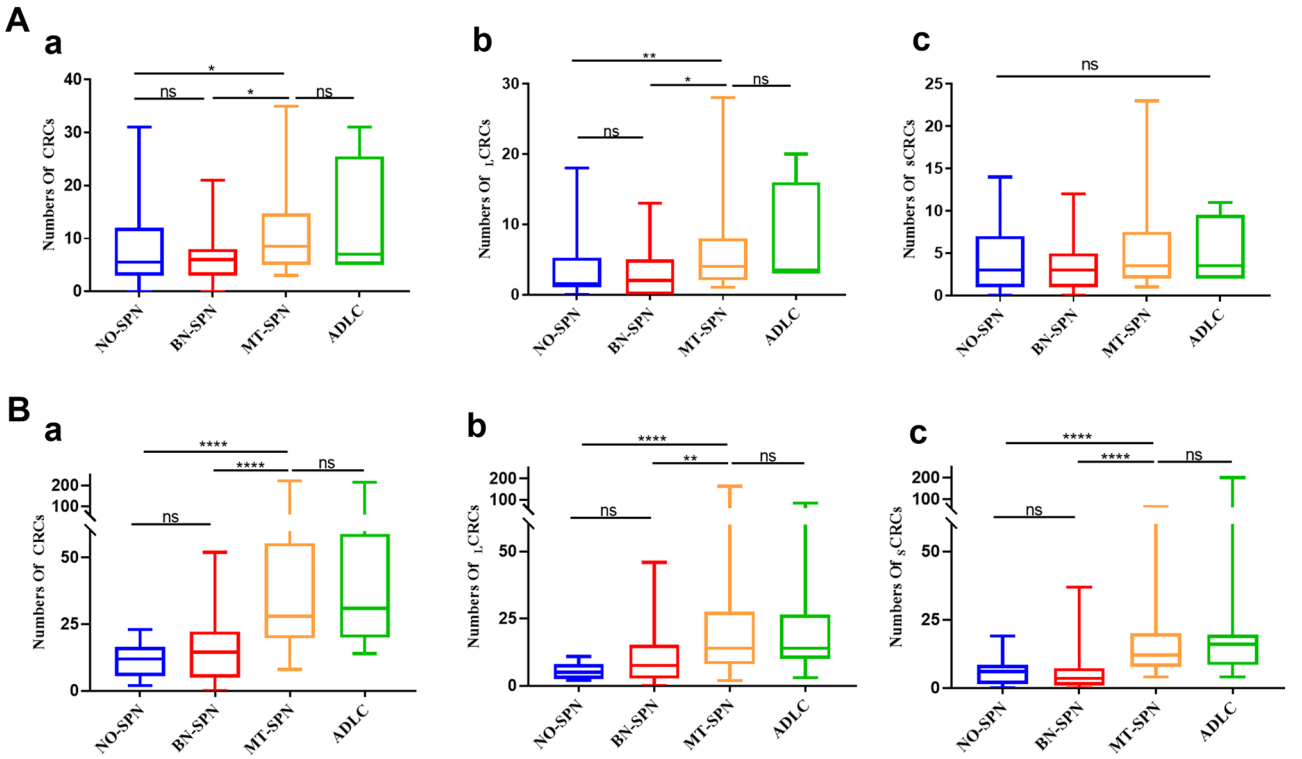


Fig. 2 The comparison of CRCs prescreening by manual observation and Metafer-image scanning assay. **A(a–c)** Comparison of the numbers of CRCs in different sizes by manual observation between groups. (non-operated solitary pulmonary nodule (NO-SPN), benign solitary pulmonary nodule (BN-SPN), malignant solitary pulmonary nodule (MT-SPN) and advanced lung cancer (ADLC)). **B(a–c)** Comparison of the numbers of CRCs in different sizes by Metafer-image scanning counting between groups. (health, non-operated solitary pulmonary nodule (NO-SPN), benign solitary pulmonary nodule (BN-SPN), malignant solitary pulmonary nodule (MT-SPN) and advanced lung cancer (ADLC)). Counting data are presented as the median, $*P<0.05$, $**P<0.01$, $***P<0.001$, $****P<0.0001$, Mann–Whitney *U* test. Blank means no statistical significance

no significant difference was observed in the CRC number between the malignant pulmonary nodule and the advanced lung cancer groups, as well as the non-operated pulmonary nodule and benign pulmonary nodule groups (Fig. 2A and B). However, when both methods were used to stratify the CRC numbers with different cell sizes, we found that the manual method was not able to show the differences in small CRCs between the malignant pulmonary nodule and benign pulmonary nodule groups (Fig. 2A-c, B-c). These results indicated that the Metafer-image scanning assay was superior to manual observation in identifying the small-size CRCs.

Comprehensive analysis of CTC and CTEC in patients with SPNs

As identification of CD31 positive aneuploidy cells in the blood sample, CRCs were considered as consisting of at least two types of cells i.e. CTCs and CTECs (Lin et al. 2017). To clarify the relationship of CTCs and CTECs with malignant SPNs, we analyzed the numbers of CTCs and CTECs in each group in more detail. Blood samples that were only detected by the Metafer-Image Scanning System were included.

The CTCs were detected in all malignant SPN and advanced lung cancer patients i.e. 100% detection rate, while the detection rates of CTCs were 82.4% (14/17), 83.3% (25/30), and 95.8% (23/24) in the non-operated SPN, benign SPN and healthy groups, respectively. CTCs with CK18+ were detected in two patients from the malignant SPN group (5.9%, 2/34) and one patient from the advanced lung cancer group (4.7% 1/21). Interestedly, the patient in the advanced lung cancer group was observed to have 118 CK18+ CTCs. The patients from the rest groups did not have any tumor marker-positive CTCs. Taken together, the iFISH showed that tumor marker positive CTCs were relatively low. The CD31 positive CTECs were also detected in all patients with malignant SPNs and advanced lung cancer (100%). The detection rates for CTECs were 94.1% (16/17), 93.3% (28/30), and 87.5% (21/24) in the non-operated SPN, benign SPN, and healthy groups, respectively. Two PD-L1+ CTECs were observed in one patient with advanced lung cancer. The CTMs were observed in the 3 patients with malignant SPN and 1 patient with advanced lung cancer,

respectively. However, the CTMs had not been observed in other groups. These results suggested that CK18+CTCs and CTMs were more prone to present in blood from patients with lung malignancies and could be critical indicators for predicting the presence of malignant lesions, which is consistent with the results reported by Mascalchi et al. (2017).

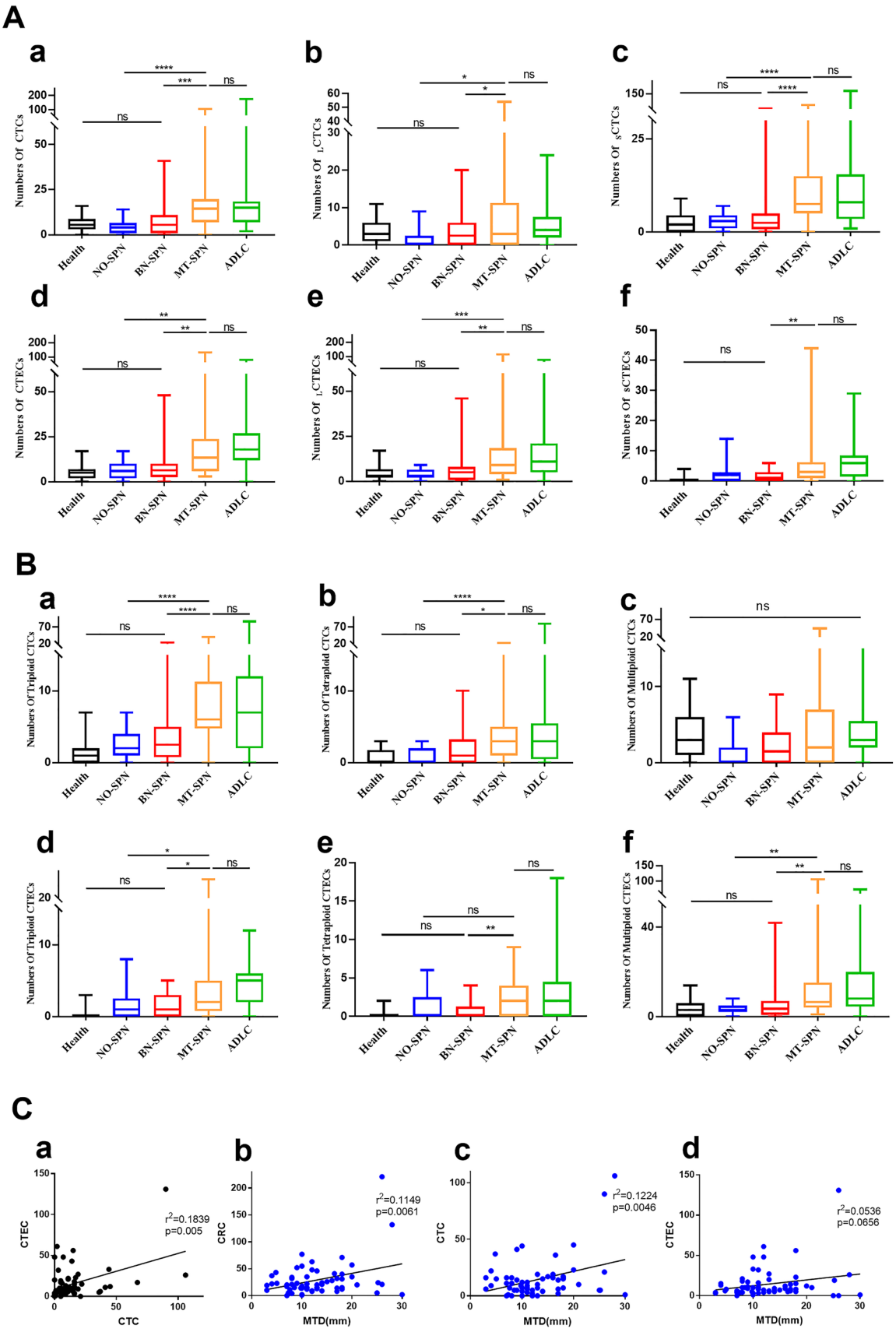
In addition, the total numbers of CTECs or CTCs in blood samples obtained from patients in the malignant SPN group were significantly higher in comparison to that in the group with benign SPNs ($P < 0.001$, $P < 0.01$, respectively, Fig. 3A-a and A-d). Interestedly, the numbers of CTECs or CTCs (no matter large or small size) in individuals with malignant SPNs increased significantly in comparison to individuals who were diagnosed as benign SPNs ($_{L}$CTCs: $P < 0.05$; $_{S}$CTCs: $P < 0.0001$; $_{L}$CTECs: $P < 0.01$; $_{S}$CTECs: $P < 0.01$, Fig. 3A-b, A-c, A-e, and A-f, respectively). However, the numbers of CTECs and CTCs were not significantly different among the non-operated SPN, the benign SPNs patients, and the healthy controls. In addition, the total number or size range of the CTECs and CTCs were similar in the malignant SPN and advanced lung cancer groups (Fig. 3A). Furthermore, the chromosomal ploidy numbers in the CTCs and CTECs were assessed among the various groups. The chromosomal aneuploidy was classified into three subtypes, including triploid, tetraploid, and multiploid (CEP 8 signal ≥ 5). The analysis revealed that the CTCs and the CTECs were often associated with the three subtypes of aneuploidy, and the number of aneuploidy CTCs and the CTECs were significantly higher in the group with malignant SPNs compared to the group with benign SPNs, except for the multiploid CTCs (Fig. 3B). Compared with the healthy controls, the differences were not significant for the aneuploidy CTCs and CTECs subtypes in the non-operated SPN group and benign SPN group. In general, the CTC and CTEC total numbers or the different aneuploidy CTCs and CTECs subtypes were similar in the malignant SPN and advanced lung cancer groups (Fig. 3B). Otherwise, we observed a significant correlation for CTCs and CTECs ($r^2 = 0.1839$, $P = 0.0050$, Fig. 3C-a).

We further analyzed whether CRCs numbers correlated with radiographically measured tumor maximum diameters. In this analysis, 64 patients whose CRC was counted by the Metafer-image scanning assay were included. The maximum tumor diameter measured by the software of CT/PET-CT ranged from 7 to 30 mm, with a mean diameter of 15.5 mm in the malignant SPN group. In contrast, the mean maximum tumor diameter in the benign SPN group was 12.3 mm (range 3–30 mm), which was significantly smaller in comparison with that of the malignant SPN group ($P = 0.012$, two-tailed Student's *t*-test). The results revealed that the numbers of either the CRCs or CTCs in patients with SPNs were indeed correlated with maximum tumor diameter shown on CT/PET-CT imaging (CRC: $r^2 = 0.1149$, $P = 0.0061$; CTC: $r^2 = 0.1224$, $P = 0.0046$; Fig. 3C-b and C–c). But no significant correlation was observed between the numbers of CTECs and the maximum tumor diameter (CTEC: $r^2 = 0.0536$, $P = 0.0656$; Fig. 3C-d).

ROC curves analysis for determining the diagnostic performance of CRCs

In comparison to the benign SPN group, ROC curves were completed using the results obtained from the malignant SPN group, including the CRC total numbers and CRC subtypes (Fig. 4A-C). Parameters with the area under ROC curve (AUC) > 0.7 were presented as follows: sCRCs (0.849) > triploid CRCs (0.841) > triploid CTCs (0.818) > sCTCs (0.800) > Total CRCs (0.778) > tetraploid CRCs (0.731) > Total CTECs (0.728) > sCTECs (0.719). Other statistical parameters of ROC analysis were shown in Table 2. The critical value that corresponded to the highest value associated with the Jordan index was determined, which should be used as cut-off values (see Table 2). The specific cut-off values were also the standards used to determine the associated specificity, sensitivity, and positive likelihood ratio. To test whether CRCs detection could be used as a non-invasive method for differentiation between malignant and benign SPNs, we compared three of the most commonly used clinically NSCLC serum tumor markers i.e. SCC, CEA, and CYFRA21-1 (Wu et al. 2020) with CRCs and their relationship with results from pathologic diagnosis. Our results suggested that CRCs detection was more reliable than serological tumor marker detection (Table 2).

In more detail, our results revealed that the sensitivity and specificity of the CRC totals in the diagnostics for malignant SPNs were 67.6% and 73.3%, with a cut-off value of 21.5 units. In addition, the specificity and sensitivity of the CTC and CTEC totals for malignant SPNs were 82.4% and 56.7%, 55.9%, and 83.3%, with cut-off values of 6.5 units and 11.5 units, respectively. Among the CRC subpopulations that exhibited an area under ROC curve (AUC) > 0.7, the small CRCs and total CTCs exhibited the higher sensitivity i.e. 82.4% with respect to the prediction of malignant SPNs, as shown in Table 2. In comparison with the routine clinical serum tumor markers and CT/PET-CT imaging, the examination of CRCs showed notable advantages with respect to differentiating between benign and malignant SPNs. In particular, the positive likelihood ratio (LR+) of triploid CRCs reached 8.4, which indicated that triploid CRCs could be a superior indicator during screening for individuals at a higher risk for lung cancer and in the early stages of the disease. This population would include individuals who smoke, have chronic lung disease, or have a family history of lung cancer. Considering that 8 mm nodule diameter is recommended as the criterion for surgical biopsy (Detterbeck et al. 2013; Wood et al. 2018), we also performed


A
a
Numbers Of CTCs
b
Numbers Of $_L$CTCs
c
Numbers Of $_S$CTCs
d
Numbers Of CTECs
e
Numbers Of $_L$CTECs
f
Numbers Of sCTECs
B
a
Numbers Of Triploid CTCs
b
Numbers Of Tetraploid CTCs
c
Numbers Of Multiploid CTCs
d
Numbers Of Triploid CTECs
e
Numbers Of Tetraploid CTECs
f
Numbers Of Multiploid CTECs
Health
NO-SPN
BN-SPN
MT-SPN
ADLC
ns
C
a
CTEC
CTC
r^2=0.1839
p=0.005
b
CRC
MTD(mm)
r^2=0.1149
p=0.0061
c
CTC
MTD(mm)
r^2=0.1224
p=0.0046
d
CTEC
MTD(mm)
r^2=0.0536
p=0.0656

◂**Fig. 3** Comprehensive analysis of CTCs and CTECs in peripheral blood between groups. **A(a–f)** Comparison of the numbers of CTCs and CTECs in different sizes between groups. (health, non-operated solitary pulmonary nodule (NO-SPN), benign solitary pulmonary nodule (BN-SPN), malignant solitary pulmonary nodule (MT-SPN) and advanced lung cancer (ADLC)). **B(a–f)** Quantification analysis of chromosomal ploidy in CTCs and CTECs between groups. (health, non-operated solitary pulmonary nodule (NO-SPN), benign solitary pulmonary nodule (BN-SPN), malignant solitary pulmonary nodule (MT-SPN) and advanced lung cancer (ADLC)). **C-a** Correlation analysis between the number of CTCs and CTECs. **C(b–d)** Correlation analysis between the number of different subtypes of CRCs and maximum tumor diameter. Count data are presented as the median, $*P<0.05$, $**P<0.01$, $***P<0.001$, $****P<0.0001$. Mann–Whitney *U* test. Blank means no statistical significance

ROC analysis on patients with pulmonary nodules < 8 mm and ≥ 8 mm. Unfortunately, due to the small sample size for pulmonary nodules < 8 mm, an effective analysis could not be performed. As to the individuals with pulmonary nodules ≥ 8 mm, the sensitivity and specificity were not significantly improved and the cut-off values of 21.5 CRC units, 6.5 CTC units, and 11.5 CTEC units were as same as the results mentioned above (data not shown).

Discussion

Solitary pulmonary nodules (SPNs) are becoming more common with the use of low-dose CT scans for screening of lung cancer. However, distinguishing malignancy from benign SPNs is a big challenge for radiologists and surgeons. The liquid biopsy is gradually mature and plays an important role in precision medicine. The detection of circulating tumor cells, one method of liquid biopsy, has been reported to be useful for diagnosis, evaluation of treatment, and prediction of prognosis. Although various techniques have been developed to detect CTCs, their advantages or limitations have also been observed. The CellSearch® system has been approved by FDA to detect CTCs in breast cancer patients which identify EpCAM positive cells in blood as CTCs. Therefore, the number set for prognosis prediction of breast cancer is low and the cutoff values are CTC ≥ 5 units/7.5 ml for the advanced-stage and CTC ≥ 1 units/7.5 ml for the early-stage breast cancer, which has been published in the updated AJCC breast cancer Staging Manual (see https://cancerstaging.org/About/news/Pages/Updated-Breast-Chapter-for-8th-Edition). The CellSearch® system has also been tried in other types of cancer but the results are not as good as breast cancer. The most critical limitation associated with CellSearch® system is that it does not take intrinsic biological characteristics such as the considerable heterogeneity of EpCAM expression on tumor cells into account. Reports have indicated that different types of tumors express different levels of EpCAM. For example, the expression of EpCAM has been reported to be low in the pancreas, bladder, and lung (Gires and Stoecklein 2014), not even observed in melanomas (Mikolajczyk et al. 2011). Bednarz-Knoll et al. (2012) reported that the down-regulation of the epithelial characteristics and up-regulation of the stem-like characteristics in NSCLC enhanced the mobility and invasiveness of epithelium-derived tumor cells, which facilitated their entry into the blood circulation and the colonization of distant organs to form metastases.

Other methods are based on the size or plasticity of the tumor cells. Such techniques include isolation based on epithelial tumor cell size ISET® (Rare Cell Diagnostics, France), Parsortix® (Angle Plc, UK), and Clear Cell® (Clearbridge BioMedics Pte Ltd, Singapore). The primary advantage of these techniques is that they assess cells independent of antigens on tumor cell surface meanwhile maintaining the activity and morphology of the isolated tumor cells. However, cells of small size can be lost easily (Alunni-Fabbroni and Sandri 2010).

In this study, we applied SE-iFISH to detect CRCs. SE-iFISH has the following advantages: 1, effective enrichment of CRCs no matter the cell size; 2, various antibodies for different tumor markers as long as you wish to identify; 3, genetic characteristics such as aneuploidy for identifying CTCs. To this day, we have found AFP, CA-199, Her-2, CD31, and PD-L1 positive cells in patients carrying different types of tumor, which are unable to be identified using the CellSearch® system or other techniques (unpublished data). It is worth to mention CD31 positive cells are the most common CRCs in our study and they are recognized as endothelial cells shed from tumor vessels or fusion cells of tumor cells with endothelial cells. Hence, its origin and clinical significance are still under investigation. In addition, whether the identification of PD-L1 positive tumor cells can help immunotherapy needs further research. Taking together, SE-iFISH is able to show tumor markers, cell size, and aneuploid at the same time and give a real-time picture of what a single CTC or CTEC looks like. The SE-iFISH technique is a more reliable platform for CRCs detection and can be expanded with the development of new antibodies or probes.

In this study, a DNA fragment that is compatible with the sequence of chromosome 8 centromere (CEP8) has been used as a probe for the detection of aneuploidy. Other probes for other chromosome centromeres can also be used and this may improve the rate of aneuploidy detection. In fact, in our follow-up studies, we are trying to use probes for chromosome 8 centromere and chromosome 12 centromere to detect aneuploidy. The primary experimental results indicated that the combination of two probes found more aneuploid cells with the same or different 8 and 12 chromosomes in a single cell which provided more valuable evidence for tumor heterogenous (unpublished data). It has been reported that chromosomal aneuploidy

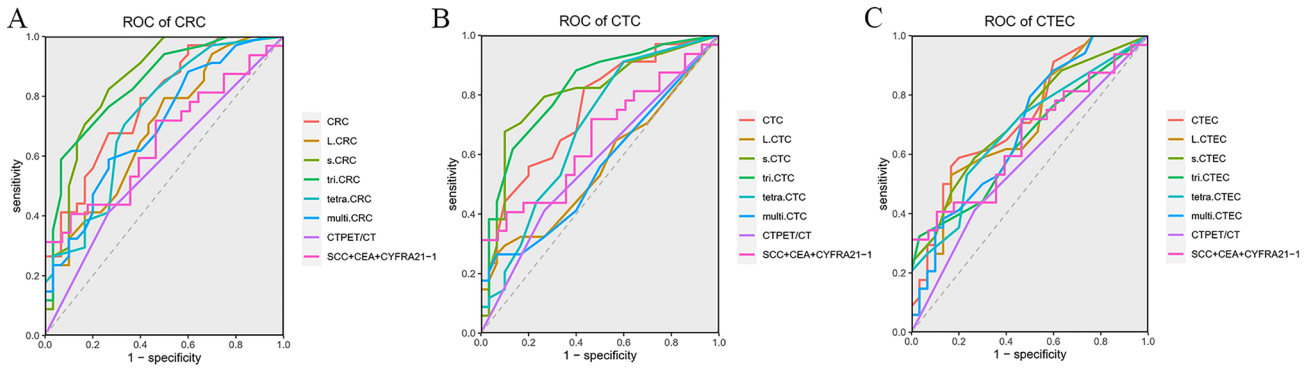


Fig. 4 ROC curve analysis of CRCs, serum tumor markers and CT/PET-CT for malignant solitary pulmonary nodules. **A** ROC curves of different subtypes of CRCs for malignant solitary pulmonary nodules. **B** ROC curves of different subtypes of CTCs for malignant solitary pulmonary nodules. **C** ROC curves of different subtypes of CTECs for malignant solitary pulmonary nodules

Table 2 Evaluation of the diagnostic efficiency of single and combined tests for SPNs

Test item	Cutoff value	AUC (95%CI)	*P*	SEN%	SPE%	LR+
Total CRCs	21.5	0.778 (0.666–0.889)	0.0001	67.6	73.3	2.5
$_{L}$CRCs	7.5	0.689 (0.560–0.818)	0.0096	79.4	50.0	1.6
$_{S}$CRCs	6.5	0.849 (0.751–0.947)	< 0.0001	82.4	73.3	3.1
Triploid CRCs	9.5	0.841 (0.745–0.938)	< 0.0001	55.9	93.3	8.4
Tetraploid CRCs	3.5	0.731 (0.607–0.855)	0.0015	70.6	66.7	2.1
Multiploid CRCs	9.5	0.699 (0.571–0.826)	0.0064	58.8	73.3	2.2
Total CTCs	6.5	0.748 (0.629–0.867)	0.0007	82.4	56.7	1.9
$_{L}$CTCs	10.5	0.559 (0.417–0.700)	0.420	26.5	93.3	4.0
$_{S}$CTCs	6.5	0.800 (0.687–0.913)	< 0.0001	67.6	90.0	6.8
Triploid CTCs	5.5	0.818 (0.713–0.923)	< 0.0001	61.8	86.7	4.6
Tetraploid CTCs	0.5	0.683 (0.550–0.815)	0.0121	91.2	40.0	1.5
Multiploid CTCs	6.5	0.557 (0.416–0.698)	0.435	26.5	93.3	4.0
Total CTECs	11.5	0.728 (0.605–0.851)	0.0018	55.9	83.3	3.4
$_{L}$CTECs	8.5	0.699 (0.569–0.828)	0.0064	52.9	83.3	3.2
$_{S}$CTECs	2.5	0.719 (0.595–0.843)	0.0027	58.8	73.3	2.2
Triploid CTECs	4.0	0.650 (0.516–0.783)	0.040	32.4	96.7	9.7
Tetraploid CTECs	1.5	0.682 (0.551–0.812)	0.013	52.9	76.7	2.3
Multiploid CTECs	3.5	0.688 (0.558–0.819)	0.0098	79.4	50.0	1.6
SCC+CEA+CYFRA 21-1	2.5,5.0,3.3	0.652 (0.513–0.790)	0.0439	50.0	71.4	1.8
CT/PET-CT Diagnosis	–	0.585 (0.444–0.727)	0.247	42.4	75.9	1.8

$_{L}$CRCs, large CRCs; $_{S}$CRCs, small CRCs; $_{L}$CTCs, large CTCs; $_{S}$CTCs, small CTCs; $_{L}$CTECs, large CTECs; $_{S}$CTECs, small CTECs; SEN, sensitivity; SPE, specificity; LR+, positive likelihood ratio

is a common biological characteristic of tumor cells and approximately 90% of solid tumors and 75% of hematologic cancers exhibit chromosome aneuploidy (Lin 2018). The majority of aneuploidy cells in blood samples showed a naked nucleus without expression of common tumor markers such as CK18 or EpCAM. Whether it implies that these aneuploidy cells might express other tumor markers which have not been identified yet or they were in the transition from epithelium to mesenchyma need to be further investigated. Nevertheless, it encourages us to explore new tumor-associated antigens.

One interesting question raised in this study is why CTCs and CTECs existed in the blood of benign SPNs patients and even healthy controls. We know that mutation happens at a different frequency in normal cells in the individual of different ages. Surprisingly we observed a certain number of CTCs and CTECs existed in healthy youths. Marquette et al. (2020) also found that CTCs existed in the population with chronic obstructive pulmonary disease (COPD) who were at the risk of lung cancer. The fact that the healthy youths and the benign SPNs patients have CTCs and CTECs suggested that abnormal aneuploidy cells occur frequently. Lei et al.

(2020) indicated that sporadic endothelial cells that were aneuploid, CD31+ and found in the blood circulatory system of healthy individuals, were due to trans-differentiation of stromal cells or were aging cells, neither of which are related to the presence of a tumor and are removed via the immune system. Thus, the complete immune function is very important for monitoring and removing abnormal aneuploidy cells and preventing tumorigenesis. Significantly, the numbers of CTCs and CTECs are evaluated in malignant SPNs than benign SPNs.

Conclusions

In summary, SE-iFISH plus auto-scan and Metafer software provided a platform for CRCs detection. With the development of more antibodies and probes, the detection of CRCs will find more types of tumor-associated CRCs and provide more valuable information for differential diagnosis not only for SPNs but also for other types of early-stage tumors. In addition, the results from our study suggested that detection of CRCs was helpful for differentiation between benign and malignant SPNs and CRCs were more informative and reliable than commonly used serum tumor markers.

Acknowledgements We thank Dr. Peter. P. Lin and his team (Cytelligen, San Diego, CA, USA) for their continuous help during the whole study.

Author contributions Dr. Q. Huang and Dr. J. Cheng contributed to the conception of the study. J. Xie, Y. Wang and B. Hu performed the experiments. J. Xie contributed significantly to data analysis and manuscript preparation. Dr. Z. Ruan and Dr. J. Zheng assisted with the patients' recruitment and sample collection. Dr. Y. Gong helped to revise the manuscript.

Funding This study was financially supported by the Shanghai "Rising Stars of Medical Talents" Youth Development Program (No. SHWRS (2021)_099 to J. Cheng).

Availability of data and materials All data generated during this study are available from the corresponding author by request.

Declarations

Conflict of interest We declare that we have no financial interests or personal relationships with other people or organizations that can inappropriately influence our work in this paper.

Ethics approval The study was performed in accordance with the Declaration of Helsinki and approved by the Ethical Review Board of Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, China (2016KY130).

Informed consent Informed consent was provided to all individuals in this study before blood collection.

Consent for publication No personal data was applicable in this manuscript.

References

Alunni-Fabbroni M, Sandri MT (2010) Circulating tumour cells in clinical practice: methods of detection and possible characterization. Methods 50(4):289–297

Bednarz-Knoll N, Alix-Panabieres C, Pantel K (2012) Plasticity of disseminating cancer cells in patients with epithelial malignancies. Cancer Metastasis Rev 31(3–4):673–687

Bidard FC, Proudhon C, Pierga JY (2016) Circulating tumor cells in breast cancer. Mol Oncol 10(3):418–430

De Luca A et al (2000) Detection of circulating tumor cells in carcinoma patients by a novel epidermal growth factor receptor reverse transcription-PCR assay. Clin Cancer Res 6(4):1439–1444

Detterbeck FC, Lewis SZ, Diekemper R, Addrizzo-Harris D, Alberts WM (2013) Executive summary: diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest 143(5 Suppl):7S-37S

Gires O, Stoecklein NH (2014) Dynamic EpCAM expression on circulating and disseminating tumor cells: causes and consequences. Cell Mol Life Sci 71(22):4393–4402

Gomez-Saez N et al (2014) Prevalence and variables associated with solitary pulmonary nodules in a routine clinic-based population: a cross-sectional study. Eur Radiol 24(9):2174–2182

Grover PK, Cummins AG, Price TJ, Roberts-Thomson IC, Hardingham JE (2014) Circulating tumour cells: the evolving concept and the inadequacy of their enrichment by EpCAM-based methodology for basic and clinical cancer research. Ann Oncol 25(8):1506–1516

Guo Qiaomei QL, Lin W, Jiatao L (2016) Diagnostic value of human circulating tumor cell detection for non-small cell lung cancer. Chin J Lab Med 39(08):589–594

Hammerschlag G, Cao J, Gumm K, Irving L, Steinfort D (2015) Prevalence of incidental pulmonary nodules on computed tomography of the thorax in trauma patients. Intern Med J 45(6):630–633

Harris L et al (2007) American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. J Clin Oncol 25(33):5287–5312

Henschke CI et al (2002) CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. AJR Am J Roentgenol 178(5):1053–1057

Ju SM et al (2013) Prevalence of non-calcified pulmonary nodules in screening chest computed tomography. Thorac Cancer 4(4):405–409

Kalluri R (2009) EMT: when epithelial cells decide to become mesenchymal-like cells. J Clin Invest 119(6):1417–1419

Kasmann L, Bolm L, Janssen S, Rades D (2017) Prognostic factors and treatment of early-stage small-cell lung cancer. Anticancer Res 37(3):1535–1537

Lei Y et al (2020) Combined detection of aneuploid circulating tumor-derived endothelial cells and circulating tumor cells may improve diagnosis of early stage non-small-cell lung cancer. Clin Transl Med 10(3):e128

Li J et al (2020) Evaluation of sensitivity and specificity of CanPatrol technology for detection of circulating tumor cells in patients with non-small cell lung cancer. BMC Pulm Med 20(1):274

Lin PP (2018) Aneuploid CTC and CEC. Diagnostics (basel) 8:2

Lin PP, Gires O, Wang DD, Li L, Wang H (2017) Comprehensive in situ co-detection of aneuploid circulating endothelial and tumor cells. Sci Rep 7(1):9789

Marquette CH et al (2020) Circulating tumour cells as a potential biomarker for lung cancer screening: a prospective cohort study. Lancet Respir Med 8(7):709–716

Mascalchi M et al (2017) Circulating tumor cells and microemboli can differentiate malignant and benign pulmonary lesions. J Cancer 8(12):2223–2230

McWilliams A et al (2013) Probability of cancer in pulmonary nodules detected on first screening CT. N Engl J Med 369(10):910–919

Mikolajczyk SD et al (2011) Detection of EpCAM-negative and cytokeratin-negative circulating tumor cells in peripheral blood. J Oncol 2011:252361

National Lung Screening Trial Research T et al (2011) Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 365(5):395–409

Ramalingam N, Jeffrey SS (2018) Future of liquid biopsies with growing technological and bioinformatics studies: opportunities and challenges in discovering tumor heterogeneity with single-cell level analysis. Cancer J 24(2):104–108

Revelo AE et al (2019) Liquid biopsy for lung cancers: an update on recent developments. Ann Transl Med 7(15):349

Tartarone A et al (2017) Possible applications of circulating tumor cells in patients with non-small cell lung cancer. Lung Cancer 107:59–64

Wang SQ et al (2020) Detection of CTCs and CSCs in the staging and metastasis of non-small cell lung cancer based on microfluidic chip and the diagnostic significance. Eur Rev Med Pharmacol Sci 24(18):9487–9496

Went PT et al (2004) Frequent EpCam protein expression in human carcinomas. Hum Pathol 35(1):122–128

Wood DE et al (2018) Lung cancer screening, Version 3.2018, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 16(4):412–441

Wu H et al (2020) The serum tumor markers in combination for clinical diagnosis of lung cancer. Clin Lab 66:3

Yu M, Stott S, Toner M, Maheswaran S, Haber DA (2011) Circulating tumor cells: approaches to isolation and characterization. J Cell Biol 192(3):373–382

Zhang T, Armstrong AJ (2016) Clinical utility of circulating tumor cells in advanced prostate cancer. Curr Oncol Rep 18(1):3

Zhang L et al (2020) PD-L1(+) aneuploid circulating tumor endothelial cells (CTECs) exhibit resistance to the checkpoint blockade immunotherapy in advanced NSCLC patients. Cancer Lett 469:355–366

Zhou J, Dong F, Cui F, Xu R, Tang X (2017) The role of circulating tumor cells in evaluation of prognosis and treatment response in advanced non-small-cell lung cancer. Cancer Chemother Pharmacol 79(4):825–833