



Fragment Enrichment of Circulating Tumor DNA With Low-Frequency Mutations

Xiaojun Liu¹, Jidong Lang^{2†}, Shijun Li^{3†}, Yuehua Wang³, Lihong Peng¹, Weitao Wang², Yingmin Han², Cuixiao Qi², Lei Song², Shuangshuang Yang², Kaixin Zhang², Guoliang Zang², Hong Pei², Qingqing Lu², Yonggang Peng², Shuxue Xi², Weiwei Wang², Dawei Yuan², Pingping Bing⁴, Liqian Zhou^{1*} and Geng Tian^{1,2*}

OPEN ACCESS

Edited by:

Tao Huang,
Shanghai Institutes for Biological
Sciences (CAS),
China

Reviewed by:

Lu Zhang,
Hong Kong Baptist University,
Hong Kong
Bing Wang,
Anhui University of Technology,
China

*Correspondence:

Jidong Lang
langjd@geneis.cn
Liqian Zhou
zhoulq11@163.com
Geng Tian
tiang@geneis.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Bioinformatics and
Computational Biology,
a section of the journal
Frontiers in Genetics

Received: 27 November 2019

Accepted: 07 February 2020

Published: 28 February 2020

Citation:

Liu X, Lang J, Li S, Wang Y, Peng L,
Wang W, Han Y, Qi C, Song L, Yang S,
Zhang K, Zang G, Pei H, Lu Q, Peng Y,
Xi S, Wang W, Yuan D, Bing P, Zhou L
and Tian G (2020) Fragment
Enrichment of Circulating Tumor DNA
With Low-Frequency Mutations.
Front. Genet. 11:147.
doi: 10.3389/fgene.2020.00147

¹ School of Computer Science, Hunan University of Technology, Zhuzhou, China, ² Bioinformatics Department, Geneis (Beijing) Co. Ltd., Beijing, China, ³ Department of Pathology, Chifeng Municipal Hospital, Chifeng, China, ⁴ Academics Working Station, Changsha Medical University, Changsha, China

Human blood contains cell-free DNA (cfDNA), with circulating tumor-derived DNAs (ctDNAs) widely used in cancer diagnosis and treatment. However, it is still difficult to efficiently and accurately identify and distinguish specific ctDNAs from normal cfDNA in cancer patient blood samples. In this study, ctDNA fragment length distribution analysis showed that ctDNA fragments are frequently shorter than the normal cfDNAs, which is consistent with previous findings. Interestingly, the ctDNA fragment length was found to be partially associated with the mutant allele frequency, with a low mutant allele frequency (< ~0.6%) associated with a longer ctDNA fragment length when compared to normal cfDNAs. The findings of this study contribute to improving the detection of low-frequency tumor mutations.

Keywords: low-frequency tumor mutation, cell-free DNA, circulating tumor-derived DNA, fragment length enrichment, mutant allele frequency, next generation sequencing

INTRODUCTION

In modern medicine, liquid biopsies are widely used in prenatal diagnoses and cancer treatment. When utilizing a liquid biopsy, circulating cell-free DNA (cfDNA), circulating tumor cells (CTCs), or exosomes are isolated for evaluation (Bardelli and Pantel, 2017; Wan et al., 2017; Siravegna et al., 2017). Of these, circulating tumor-derived DNA (ctDNA) is widely utilized as a tumor biomarker in translational and clinical research (Diaz and Bardelli, 2014; Donaldson and Park, 2018), while fetal cfDNA obtained from maternal blood is widely used as a noninvasive method for prenatal diagnoses (Lun et al., 2008; Lo et al., 2010; Yu et al., 2014; Sun et al., 2018).

About 30 years ago, Stroun et al. first discovered that cancer patient blood samples contain cfDNA of cancer origin (Stroun et al., 1989; Thierry et al., 2016). In the following decades, ctDNA has been gradually developed as a clinical tool for cancer diagnosis and treatment, and has even been used as a prognostic or predictive factor (Mao et al., 1994; Lecomte et al., 2002; Kimura, 2006; Diehl et al., 2008). Currently, the use of ctDNA detection in cancer therapy has been approved by the US Food and Drug Administration as a treatment determinant (osimertinib or erlotinib) in non-small-cell lung carcinoma (NSCLC) patients with an *EGFR* mutation in the event that a tumor biopsy cannot be performed (US Food & Drug Administration, 2016). The application of ctDNA in cancer therapy is reliant on precise

polymerase chain reaction (PCR)-based technologies, such as droplet digital PCR (ddPCR) or amplification refractory mutation system (ARMS)-PCR, and deep-sequencing technologies; these techniques aid in distinguishing ctDNAs from other normal cfDNAs within the plasma and enable hotspot mutation detection within cancer driver genes (Taly et al., 2013; Newman et al., 2014; Frenel et al., 2015; Azizi et al., 2018). However, ctDNAs are usually present in low abundance relative to the normally occurring cfDNAs derived from normal cells, particularly in non-metastatic solid tumors (Tug et al., 2014; Siravegna et al., 2017). Consequently, there is an urgent need to reliably distinguish ctDNAs from normal cfDNA to improve the accuracy of identifying driver gene mutations.

Recently, tumor-derived ctDNAs have been shown to vary in size and are shorter than normal cfDNAs in healthy people (Umetani et al., 2006; Thierry et al., 2010; Mouliere et al., 2011; Mouliere et al., 2013). This trend was also observed during pregnancy, with fetal cfDNA usually of a different fragment size than the maternal cfDNA (Lun et al., 2008; Lo et al., 2010). Furthermore, in one study examining ctDNA length distributions in hepatocellular carcinoma (HCC) patients, copy number aberrations were leveraged and showed that high-concentration ctDNA fractions were more fragmented, while low-concentration fractions were paradoxically longer (Jiang et al., 2015; Mouliere and Rosenfeld, 2015; Jiang and Lo, 2016). In another study, ctDNAs were found to be consistently shorter than normal cfDNA, in both animal xenograft models and clinical plasma samples (Underhill et al., 2016). Additionally, mutant ctDNA fragments from tumor patients were always shorter than wild-type cfDNA fragments from healthy donors, with mutant alleles more commonly having shorter fragment lengths, something that could potentially be exploited to improve ctDNA detection (Underhill et al., 2016; Hellwig et al., 2018). Moreover, a later study confirmed that this size difference could be exploited to enhance sensitivity when monitoring ctDNAs and for noninvasive genomic analysis of various cancers (Mouliere et al., 2018). However, few studies have examined the impact of mutant allele frequency on the size distribution of ctDNA fragments, and most studies were conducted in cancer patients with relatively high mutant allele frequencies.

Thus, the aim of this study was to examine ctDNA fragment distributions in patients with low mutant allele frequencies and determine whether the ctDNA fragment length is affected by the mutant allele frequency. This was accomplished by utilizing blood samples from cancer patients with a variety of different histological types and stages. Key driver gene mutation frequencies were determined using deep-sequencing technologies and ddPCR, and fragment length differences between mutant ctDNAs and normal cfDNAs obtained from the cancer patient samples were examined.

MATERIALS AND METHODS

Sample Collection

All 105 samples (male: 49.52%, female: 50.48%) were obtained from lung cancer patients from Chifeng Municipal Hospital. All

patients provided informed written consent before de-identification. The median age of the patients was 63.5 years old (range from 36 to 85 years old). Our research was approved by the Medical Research Ethics Committee of Chifeng Municipal Hospital (Ethics [2018] No. 017).

Next-Generation Sequencing (NGS) Library Preparation, Sequencing, and Bioinformatics

Cell-free DNA was extracted using a QIAamp Circulating Nucleic Acid Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The extracted DNA (20 ng/sample) was then used to build libraries using Accel-NGS[®] 2S Plus DNA Library Kits (96 reactions; Swift BioSciences, Ann Arbor, MI, USA). Customized probes were obtained from Integrated DNA Technologies (IDT, Skokie, IL, USA) and were used for hybridization capture. All cfDNA libraries utilized a 38-hotspot gene panel (**Supplementary File**) and were quantified using a Universal Library Quantification Kit (Kapa Biosystems, Wilmington, MA, USA) on an ABI 7500 Real-Time PCR system (Applied Biosystems, Waltham, MA, USA). Sample quality was evaluated using a high sensitivity DNA kit (Agilent Technologies, Santa Clara, CA, USA) with an Agilent 2100 Bioanalyzer as per the manufacturer's instructions. NGS with fusion detection was performed using a NextSeq 500/550 High Output v2 kit with a NextSeq 500 sequencer (Illumina, San Diego, CA, USA) for 302 cycles, with standing paired-end reads of 151 bp (average sequencing depth was ~2,164X, details in the **Supplementary File**).

The FASTQ reads were collapsed into unique observations based on barcodes using CASAVA (v1.8.2) software. Low-quality and adapter-contaminated reads were removed from the raw reads using Cutadapt (v1.12) and aligned to the Hg19 reference genome using the Burrow-Wheeler Aligner for short-read alignment (bwa aln; 0.7.12-r1039). Paired-end reads with hotspots were extracted from the paired-end alignment information (column 9th) in BAM format using Samtools (v0.1.19-44428cd), and the corresponding insert size information was extracted. Finally, the extracted paired-end reads were aligned to the Hg19 reference genome again using SOAP (2.21), and hotspot mutation fragment lengths and wild fragment lengths were calculated with the alignment mismatch information (column 11th) in the alignment files.

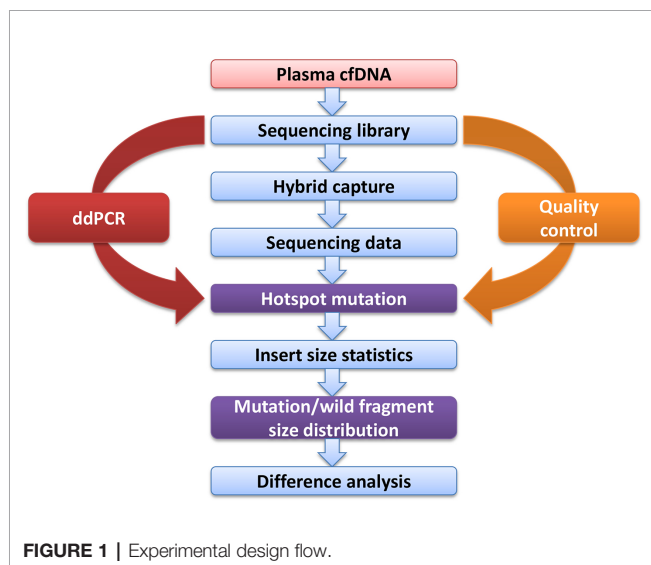
Digital Droplet PCR

EGFR-T790M, *EGFR-L858R*, *BRAF-V600E*, *PIK3CA-E545K*, *KRAS-G12C*, and *KRAS-G12V* mutant allele frequencies were determined using a Digital Droplet PCR system (Bio-Rad Laboratories, Inc., Hercules, CA, USA), with a droplet size of 1 nL in a total reaction volume of 20 μ L, with ~20 ng of cfDNA library utilized. All primers and probes were synthesized by IDT (Skokie, IL, USA; **Table 1**). Droplet counts were determined using the QuantaSoft software (Bio-Rad).

For the *PIK3CA-E545K* (n = 1), *KRAS-G12C* (n = 3), *KRAS-G12V* (n = 1), *EGFR-T790M* (n = 5), and *EGFR-L858R* (n = 1) samples, amplified libraries were utilized prior to size selection to

TABLE 1 | Primers and probes used in droplet digital PCR experiments.

| Mutation | Forward primer | Reverse primer | Wild probe | Mutation probe |
|----------|--------------------------------|-------------------------------|---------------------------------------|---------------------------------------|
| T790M | GCCTGCTGGGCATCTG | TCTTTGTGTTCCCGGACATAGTC | VIC- ATGAGCTGCGTGATGAG- MGB-NFQ | FAM- ATGAGCTGCATGATGAG- MGB-NFQ |
| L858R | GCAGCATGTCAAGATCACAGATT | CCTCCTTCTGCATGGTATTCTTTCT | VIC- AGTTTGGCCAGCCCAA- MGB-NFQ | FAM- AGTTTGGCCCGCCCAA- MGB-NFQ |
| V600E | CATGAAGACCTCACAGTAAAAATAGGTGAT | TGGGACCCACTCCATCGA | VIC- CTAGCTACAGTGAAATC- MGB-NFQ | FAM- TAGCTACAGAGAAATC- MGB-NFQ |
| E545K | CACTTACCTCTGACTCCATAGAAAATCTT | AAAGCAATTACTACAGATATCCTCTCTC | HEX-TCCTGCTCAGTGATT- MGB-NFQ | FAM-CTCCTGCTTAGTGATT- MGB-NFQ |
| G12 G12V | AATTAGATGTATCGTCAAGGCACTCTT | GCTGAAAATGACTGAATATAAACTTGTGG | VIC-TACGCCACCAGCTC- MGB-NFQ | FAM-TACGCCAACAGCTC- MGB-NFQ |
| G12C | | | | FAM-TACGCCACAAGCTCT- MGB-NFQ |



define gates for wild-type and mutation droplet populations. Libraries were constructed using the obtained DNA (~20 ng) and a Rapid DNA Lib Prep kit (ABclonal, Woburn, MA, USA). The obtained libraries (~1.2 mg) were then separated using 2% agarose gel electrophoresis and bands between 130–160 bp and 160–230 bp were extracted using a QIAquick Gel Extraction kit (Qiagen). All of the selected fragment size libraries were then validated using ddPCR as described above (**Supplementary File**).

TABLE 2 | Summary of the mutation frequencies based on next generation sequencing.

| | T790M | L858R | V600E | E545K | G12C | G12V |
|---------------------------------|-------------|-------------|------------|-------------|-------------|------------|
| Validation library number | 32 | 28 | 13 | 13 | 13 | 6 |
| Low mutation frequency 0.1–1% | 16 | 11 | 13 | 11 | 4 | 5 |
| Medium mutation frequency 1–10% | 10 | 12 | 0 | 1 | 7 | 1 |
| High-mutation frequency 10–100% | 6 | 5 | 0 | 1 | 2 | 0 |
| Mutation frequency distribution | 0.11–74.75% | 0.15–35.77% | 0.10–0.30% | 0.10–21.67% | 0.10–33.81% | 0.11–1.26% |

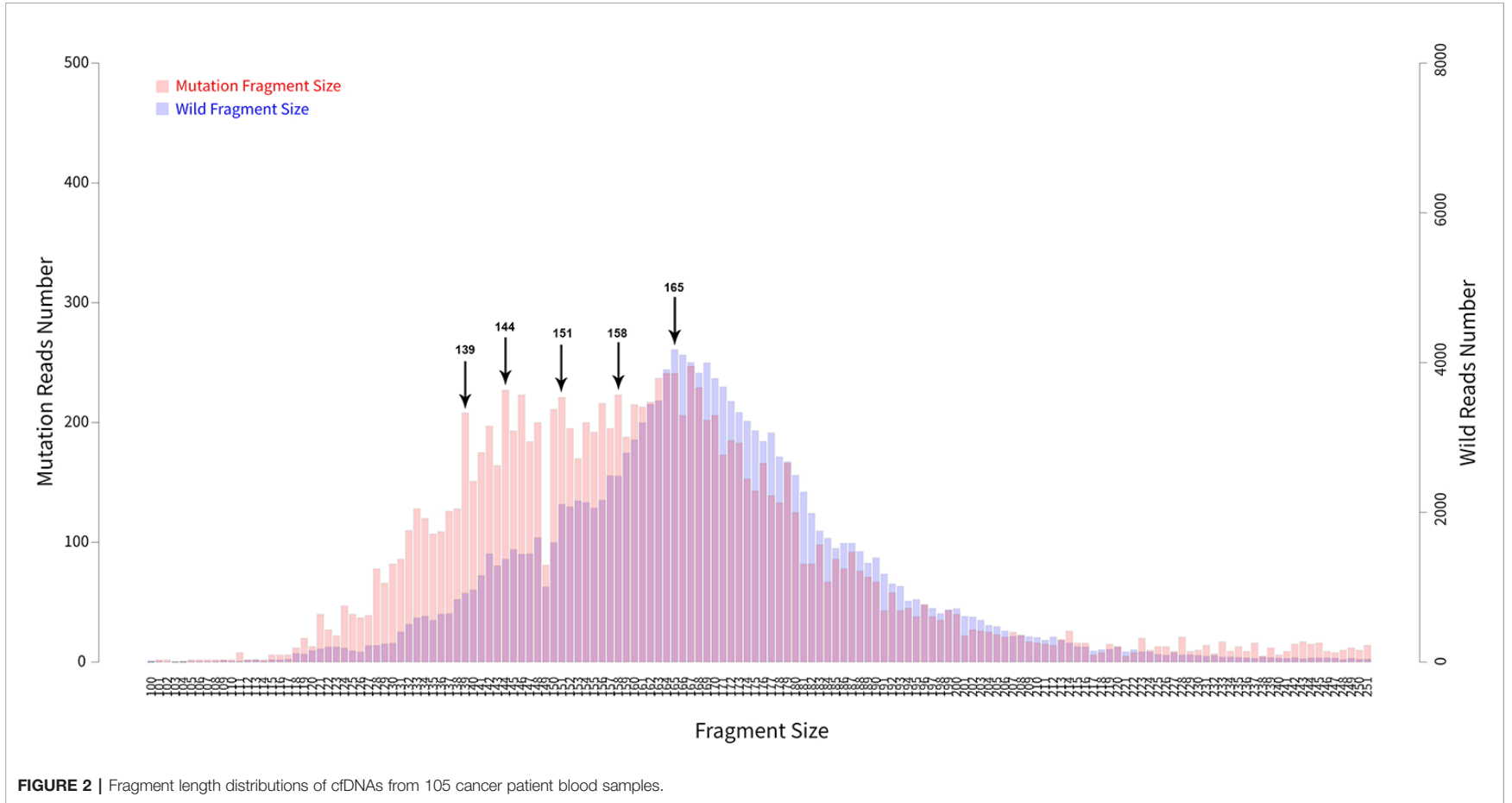
RESULTS

Comparison of cfDNA Fragment Sizes in Cancer Patient Plasma Samples

Blood samples were obtained from cancer patients with defined driver gene hotspot mutations, including *EGFR*-T790M (n = 32), *EGFR*-L858R (n = 28), *BRAF*-V600E (n = 13), *PI3KCA*-E545K (n = 13), *KRAS*-G12C (n = 13), and *KRAS*-G12V (n = 6). The cfDNA-sequencing libraries were analyzed by both NGS and ddPCR to precisely detect the mutant allele frequencies of these six hotspots in each cancer patient (**Figure 1**). Some hotspot mutant allele frequencies were more variable, such as *EGFR*-T790M (0.11–74.75%), *EGFR*-L858R (0.15–35.77%), *PI3KCA*-E545K (0.10–21.67%), and *KRAS*-G12C (0.10–33.81%; **Table 2**). Furthermore, some hotspot allele frequencies within these driver genes were relatively low, including *BRAF*-V600E (0.10–0.30%) and *KRAS*-G12V (0.11–1.26%; **Table 2**), which could be explained by examining cancer patient mutant allele frequencies, whole cfDNA fragment length distributions were globally observed. As expected, the mutant ctDNA fragments were generally shorter than the normal cfDNAs (**Figures 2 and 3**). In patients with a low mutation frequency, the ctDNA

Mutant Alleles Have a Shorter Fragment Length Than the Wild-Type Alleles

In addition to examining cancer patient mutant allele frequencies, whole cfDNA fragment length distributions were globally observed. As expected, the mutant ctDNA fragments were generally shorter than the normal cfDNAs (**Figures 2 and 3**). In patients with a low mutation frequency, the ctDNA



fragment length was longer than the normal cfDNAs, such as *BRAF*-V600E (0.10–0.30%). However, this trend was not observed in the four other DNA fragment size distribution, including *EGFR*-T790M (0.11–74.75%), *EGFR*-L858R (0.15–35.77%), *PI3KCA*-E545K (0.10–21.67%), *KRAS*-G12V (0.11–1.26%) or *KRAS*-G12C (0.10–33.81%; **Figure 3**).

Longer Fragment Lengths in Mutant ctDNAs With a Low Mutation Frequency

Fragment size differences between cancer patient ctDNAs and normal cfDNAs were further examined in conjunction with a low, medium, or high mutant allele frequency. In fragments associated with a low mutant allele frequency, the ctDNA fragments were longer than the normal cfDNAs (**Figure 4**), such as *EGFR*-T790M (0.22 and 0.21%). However, in ctDNAs with a higher mutant allele frequency, such as *EGFR*-T790M (74.75%), or a medium frequency, such as *EGFR*-T790M (4.57%), fragment lengths were shorter than the normal cfDNAs (**Figure 5** and **Table 3**).

Low-Frequency Mutations Are Associated With Large Fragment Sizes

Different fragment sizes were observed among the mutant ctDNAs, including long ctDNA (longer than normal cfDNAs), normal ctDNA (comparable to normal cfDNA lengths), and short ctDNA (shorter than normal cfDNAs). Within these three

groups, the mutant allele frequency distributions were examined and showed that a low mutation frequency was commonly associated with a long ctDNA fragment length, while normal and short ctDNAs were not (**Figure 6**).

Enrichment of Longer ctDNA Fragments Could Improve the Detection of Low-Frequency Mutations

After discovering that a low-frequency is associated with a longer ctDNA fragment size, this study aimed to determine if enriching longer ctDNA fragments could increase the mutation frequency in blood samples with a low mutant allele frequency. In one patient with a high frequency for *EGFR*-T790M (44.53%), ctDNA was extracted and different fragment sizes were obtained. To further detect the *EGFR*-T790M frequency, DNA libraries comprising two different DNA fragment sizes were examined using ddPCR. The *EGFR*-T790M frequency in a library with a fragment length between 160 and 230 bp (42.20%) was lower than the library with a fragment size between 130–160 bp (46.40%; **Figure 7A**). This was consistent with the findings presented above. Conversely, a cfDNA sample was obtained from a patient with a low *EGFR*-T790M frequency (0.54%) and different fragment sizes were collected and analyzed. In the library with fragment sizes between 160–230 bp, the *EGFR*-T790M frequency was increased (1.04%) when compared to the library with fragment sizes between 130–160 bp (0.30%; **Figure 7B**).

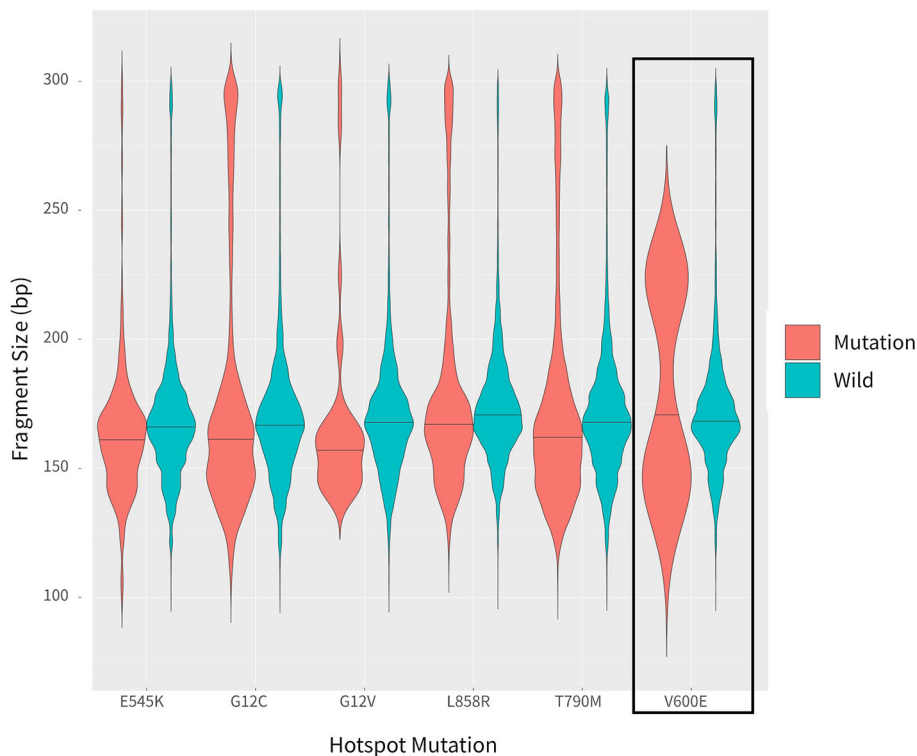


FIGURE 3 | Comparison of fragment length sizes between ctDNAs and normal cfDNAs.

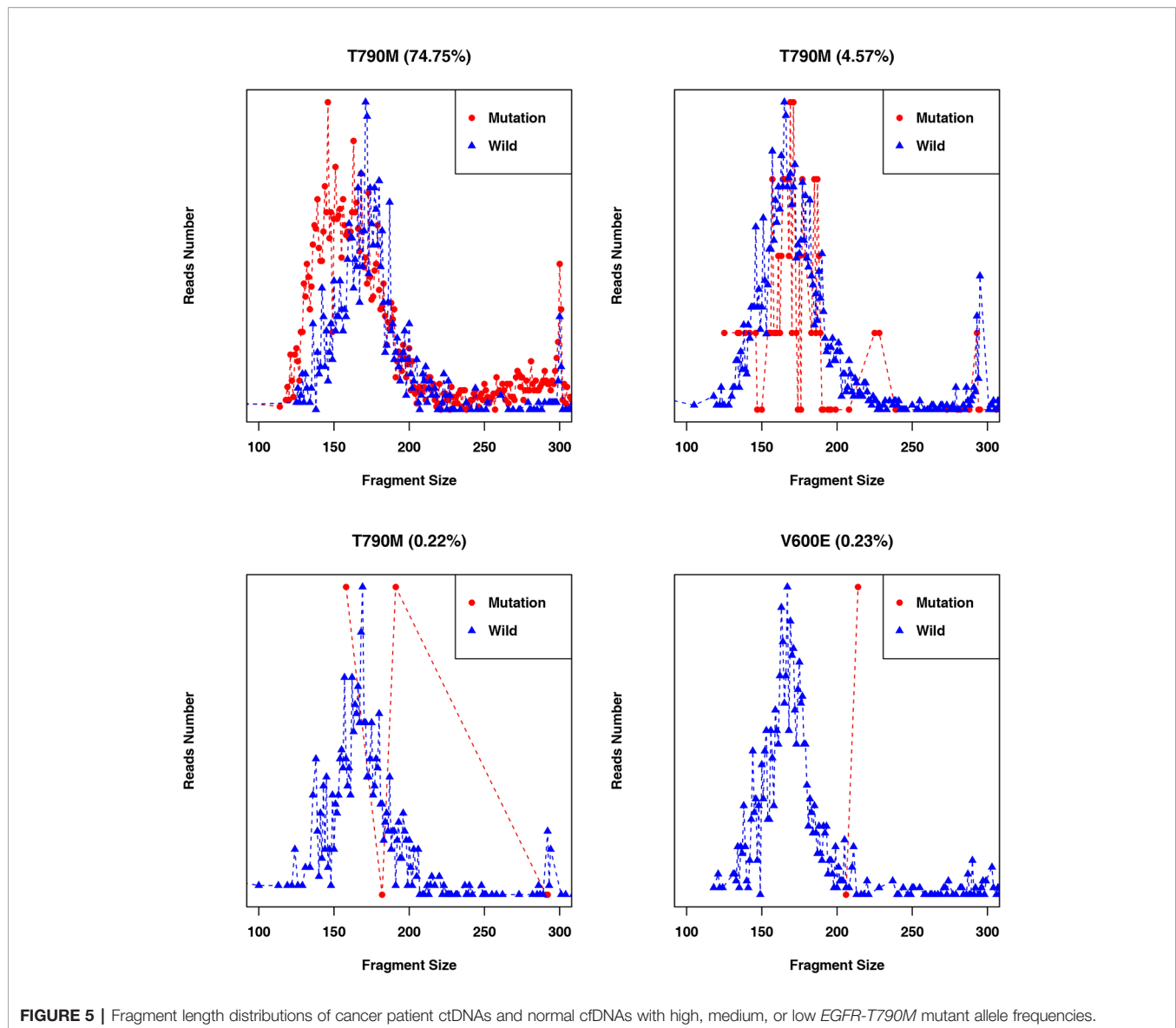


FIGURE 5 | Fragment length distributions of cancer patient ctDNAs and normal cfDNAs with high, medium, or low *EGFR-T790M* mutant allele frequencies.

TABLE 3 | Fragment length distributions of cancer patient ctDNAs and normal cfDNAs with high, medium, or low *EGFR-T790M* mutant allele frequencies.

| Mutation type | NGS (%) | ddPCR (%) | Description | Mutation peak | Wild peak | Mutation fragment median | Wild fragment median |
|---------------|---------|-----------|-------------|---------------|-----------|--------------------------|----------------------|
| <i>T790M</i> | 74.75 | 69.33 | Short | 146 | 171 | 164 | 172 |
| <i>T790M</i> | 4.57 | 5.55 | Long | 169/171 | 165 | 169.5 | 169 |
| <i>T790M</i> | 0.22 | 0.26 | Other | 158/191 | 169 | 191 | 168 |
| <i>V600E</i> | 0.23 | 0.17 | Long | 214 | 167 | 214 | 168 |

DISCUSSION

This study showed that a consistent fragment length difference occurs when comparing ctDNAs and normal cfDNA, with the mutant allele almost always associated with a shorter ctDNA fragment size, which is consistent with previous findings (Jiang et al., 2015; Underhill et al., 2016; Moulriere et al., 2018). However, some mutant ctDNAs were found to have a longer

fragment size when compared to normal cfDNAs and were associated with a low mutant allele frequency, which has not been previously reported. Furthermore, this study showed that in cancer patient plasma samples, the ctDNA fragment length is associated with the mutant allele frequency and may even be affected by it.

Here, blood samples were obtained from 105 patients that contained different cancer driver gene mutations, such as NSCLC

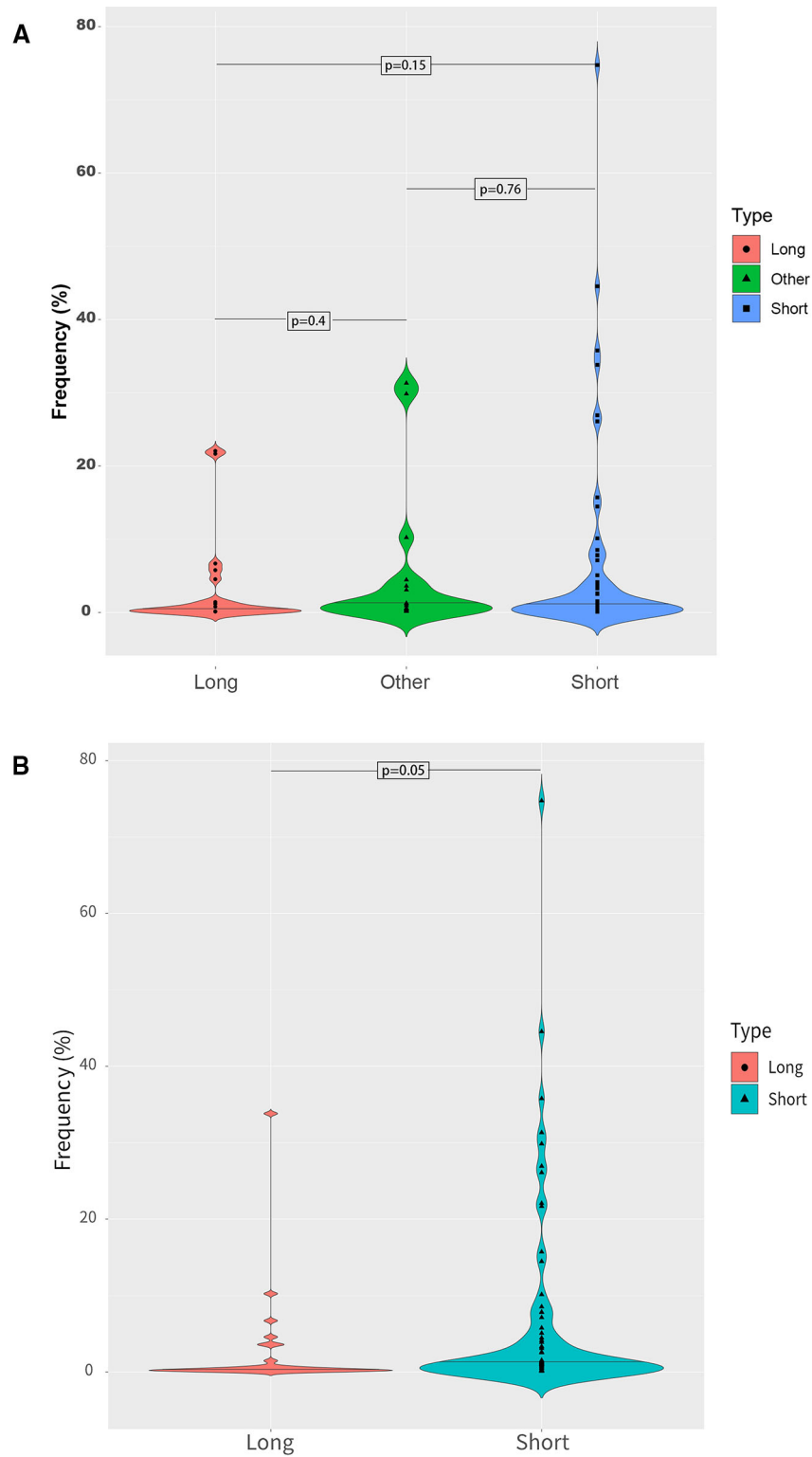


FIGURE 6 | Mutant allele frequency distributions based on ctDNA fragment length. **(A)** Relationship between the mutation fragment size peak and the mutation frequency. **(B)** Relationship between the median mutation fragment size peak and the mutation frequency.

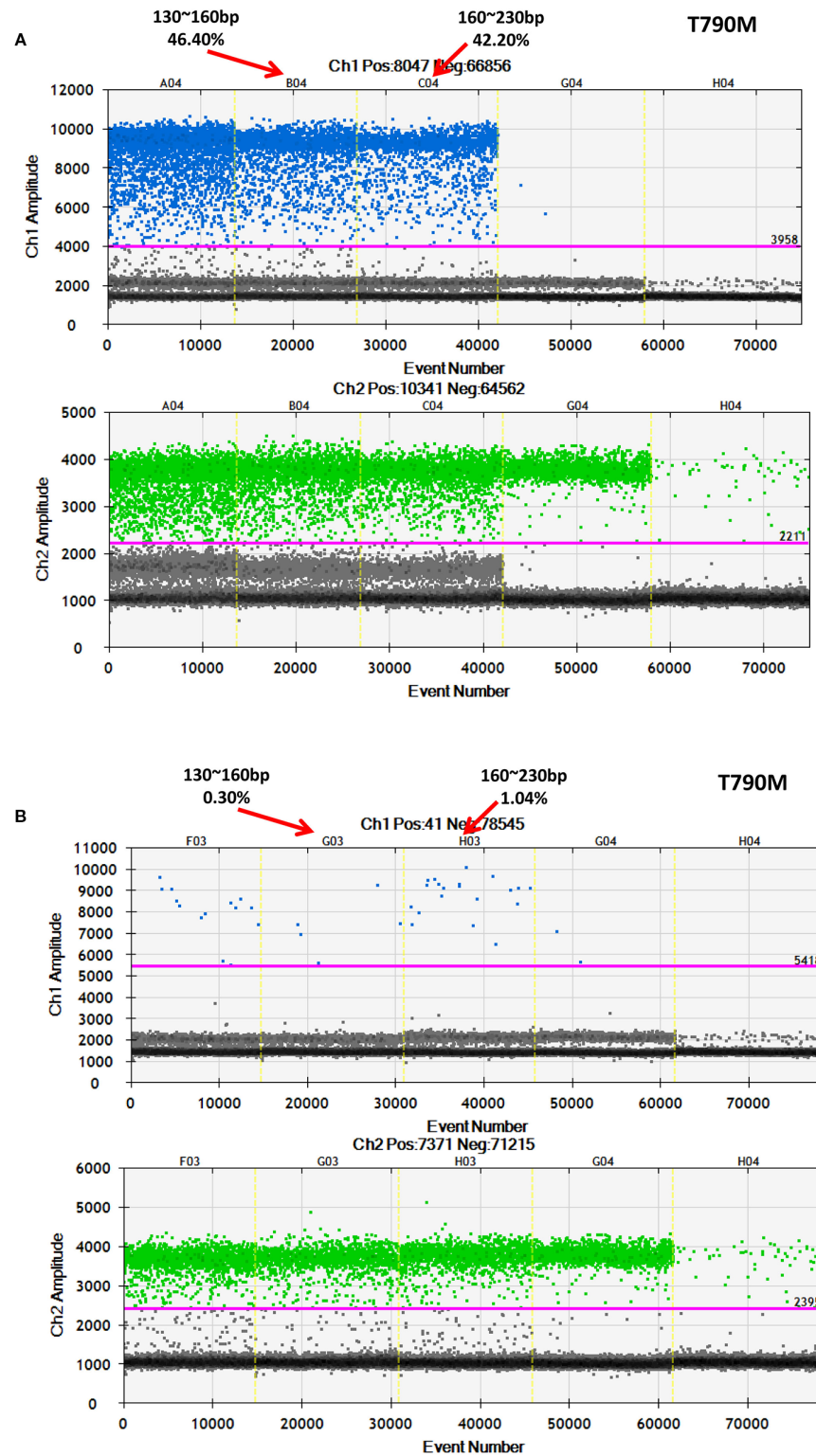


FIGURE 7 | Further validation of an association between fragment size and frequency using ddPCR. Examination of different fragment size libraries from a patient with **(A)** a high *EGFR-T790M* frequency (44.53%) and from a patient with **(B)** a low *EGFR-T790M* frequency (0.54%) using ddPCR.

patients with an *EGFR* gene mutation and colorectal cancer patients with a *BRAF* mutation. In general, mutant ctDNA fragments were found to be much shorter than normal cfDNA fragments regardless of the histological type or driver gene mutation. However, ctDNA fragments with a low mutant allele frequency were found to be longer than normal cfDNA fragments. In another study, longer mutant ctDNA fragments were also detected in cancer patient blood samples, but this phenomenon could not be explained at the time (Mouliere et al., 2018). The findings presented herein may partially explain the origin of these longer mutant ctDNA fragments.

In a previous study examining HCC plasma samples, ctDNAs with low fractional concentrations were also found to have a longer size distribution relative to the healthy controls (Jiang et al., 2015), which is similar to the observations in this study. However, the previous study only compared fragment length differences between cancer patients and healthy donors, and did not distinguish mutant ctDNA fragments from normal cfDNAs due to experimental design limitations. Taken together, these findings could suggest that early-stage tumors tend to release longer ctDNA fragments at a low-frequency, but this hypothesis requires further examination.

Mutant ctDNA fragments with a low allele frequency are hard to be accurately detected. Here, two advanced technologies to detect mutant ctDNA fragments and monitor mutant allele frequency were employed to overcome this obstacle. The cfDNA fragment sizes were accurately determined using deep-sequencing technologies, and the mutant allele frequencies were further confirmed using ddPCR. However, even these advanced technologies are susceptible to false positives.

Furthermore, the lost enrichment phenomenon of short fragments observed in this study may be related to factors such as the designed probe size (120 bp), cfDNA purification, and library construction. Moreover, the findings presented herein indicate that size selection can further improve the ctDNA detection rate and accuracy. Additionally, it would seem that when constructing a ctDNA library for early-stage cancer patients, a larger DNA fragment size (> 167 bp) should be enriched, while in later stages, enrichment of shorter DNA fragment size (< 167bp) is more beneficial.

In summary, this study demonstrates that plasma ctDNAs are generally shorter than normal cfDNAs. However, for cancer patients with a low mutant allele frequency or early tumor stage, mutant ctDNA fragments are longer than normal cfDNAs. These findings may potentially facilitate the accurate detection of cancer gene mutations when utilizing liquid biopsies, and improve the application of ctDNA detection in early cancer diagnoses.

REFERENCES

- Azizi, E., Carr, A. J., Plitas, G., Cornish, A. E., Konopacki, C., Prabhakaran, S., et al. (2018). Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell*. 174 (5), 1293–1308.e36. doi: 10.1016/j.cell.2018.05.060
- Bardelli, A., and Pantel, K. (2017). Liquid biopsies, what we do not know (Yet). *Cancer Cell* 31, 172–179. doi: 10.1016/j.ccell.2017.01.002

DATA AVAILABILITY STATEMENT

FASTQ data files for this study can be found in the NCBI Sequence Read Archive (SRA) database (BioProject ID: PRJNA562379).

ETHICS STATEMENT

Our research was approved by the Medical Research Ethics Committee of Chifeng Municipal Hospital (Ethics [2018] No. 017). All patients provided informed written consent before de-identification.

AUTHOR CONTRIBUTIONS

GT, JL, and LZ designed the project and analyzed the data. XL, WTW, and JL wrote the manuscript. SL, YW, LP, YH, SY, GZ, SX, and HP collected the data. CQ, LS, and KZ did the ddPCR experiments. WWW, DY, YP, QL, and PB modified and reviewed the manuscript.

FUNDING

This research was funded by the Natural Science Foundation of China (Grant 61803151), the Chifeng Municipal Hospital In-Situ Science and Technology Plan Project (2017-KY012), the Natural Science Foundation of Hunan province (Grant 2018JJ3570) and the Project of Scientific Research Fund of Hunan Provincial Education Department (Grant 17A052).

ACKNOWLEDGMENTS

We thank Xu Chu for help adjusting the figures.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00147/full#supplementary-material>

- Diaz, L. A., and Bardelli, A. (2014). Liquid biopsies: genotyping circulating tumor DNA. *J. Clin. Oncol.* 32, 579–586. doi: 10.1200/JCO.2012.45.2011
- Diehl, F., Schmidt, K., Choti, M. A., Romans, K., Goodman, S., Li, M., et al. (2008). Circulating mutant DNA to assess tumor dynamics. *Nat. Med.* 14, 985–990. doi: 10.1038/nm.1789
- Donaldson, J., and Park, B. H. (2018). Circulating tumor DNA: measurement and clinical utility. *Annu. Rev. Med.* 69, 223–234. doi: 10.1146/annurev-med-041316-085721

- Frenel, J. S., Carreira, S., Goodall, J., Roda, D., Perez-Lopez, R., Tunariu, N., et al. (2015). Serial next-generation sequencing of circulating cell-free DNA evaluating tumor clone response to molecularly targeted drug administration. *Clin. Cancer Res.* 21, 4586–4596. doi: 10.1158/1078-0432.CCR-15-0584
- Hellwig, S., Nix, D. A., Gligorich, K. M., O'Shea, J. M., Thomas, A., Fuentes, C. L., et al. (2018). Automated size selection for short cell-free DNA fragments enriches for circulating tumor DNA and improves error correction during next generation sequencing. *PLoS One* 13, e0197333. doi: 10.1371/journal.pone.0197333
- Jiang, P., and Lo, Y. M. D. (2016). The long and short of circulating cell-free DNA and the ins and outs of molecular diagnostics. *Trends Genet.* 32, 360–371. doi: 10.1016/j.tig.2016.03.009
- Jiang, P., Chan, C. W. M., Chan, K. C. A., Cheng, S. H., Wong, J., Wong, V. W.-S., et al. (2015). Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *Proc. Natl. Acad. Sci.* 112, E1317–E1325. doi: 10.1073/pnas.1500076112
- Kimura, H. (2006). Detection of epidermal growth factor receptor mutations in serum as a predictor of the response to gefitinib in patients with non-small-cell lung cancer. *Clin. Cancer Res.* 12, 3915–3921. doi: 10.1158/1078-0432.CCR-05-2324
- Lecomte, T., Berger, A., Zinzindohoué, F., Micard, S., Landi, B., Blons, H., et al. (2002). Detection of free-circulating tumor-associated DNA in plasma of colorectal cancer patients and its association with prognosis: plasma DNA in colorectal cancer patients. *Int. J. Cancer* 100, 542–548. doi: 10.1002/ijc.10526
- Lo, Y. M. D., Chan, K. C. A., Sun, H., Chen, E. Z., Jiang, P., Lun, F. M. F., et al. (2010). Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Sci. Transl. Med.* 2, 61ra91–61ra91. doi: 10.1126/scitranslmed.3001720
- Lun, F. M. F., Tsui, N. B. Y., Chan, K. C. A., Leung, T. Y., Lau, T. K., Charoenkwan, P., et al. (2008). Noninvasive prenatal diagnosis of monogenic diseases by digital size selection and relative mutation dosage on DNA in maternal plasma. *Proc. Natl. Acad. Sci.* 105, 19920–19925. doi: 10.1073/pnas.0810373105
- Mao, L., Hruban, R. H., Boyle, J. O., Tockman, M., and Sidransky, D. (1994). Detection of oncogene mutations in sputum precedes diagnosis of lung cancer. *Cancer Res.* 54, 1634–1637.
- Mouliere, F., and Rosenfeld, N. (2015). Circulating tumor-derived DNA is shorter than somatic DNA in plasma. *Proc. Natl. Acad. Sci.* 112, 3178–3179. doi: 10.1073/pnas.1501321112
- Mouliere, F., Robert, B., Arnau Peyrotte, E., Del Rio, M., Ychou, M., Molina, F., et al. (2011). High fragmentation characterizes tumour-derived circulating DNA. *PLoS One* 6, e23418. doi: 10.1371/journal.pone.0023418
- Mouliere, F., El Messaoudi, S., Gongora, C., Guedj, A.-S., Robert, B., Del Rio, M., et al. (2013). Circulating cell-free DNA from colorectal cancer patients may reveal high KRAS or BRAF mutation load. *Transl. Oncol.* 6, 319–328. doi: 10.1593/tlo.12445
- Mouliere, F., Chandrananda, D., Piskorz, A. M., Moore, E. K., Morris, J., Ahlborn, L. B., et al. (2018). Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci. Transl. Med.* 10, eaat4921. doi: 10.1126/scitranslmed.aat4921
- Newman, A. M., Bratman, S. V., To, J., Wynne, J. F., Eclov, N. C. W., Modlin, L. A., et al. (2014). An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat. Med.* 20, 548–554. doi: 10.1038/nm.3519
- Siravegna, G., Marsoni, S., Siena, S., and Bardelli, A. (2017). Integrating liquid biopsies into the management of cancer. *Nat. Rev. Clin. Oncol.* 14, 531–548. doi: 10.1038/nrclinonc.2017.14
- Stroun, M., Anker, P., Maurice, P., Lyautey, J., Lederrey, C., and Beljanski, M. (1989). Neoplastic characteristics of the DNA found in the plasma of cancer patients. *Oncology* 46, 318–322. doi: 10.1159/000226740
- Sun, K., Jiang, P., Wong, A. I. C., Cheng, Y. K. Y., Cheng, S. H., Zhang, H., et al. (2018). Size-tagged preferred ends in maternal plasma DNA shed light on the production mechanism and show utility in noninvasive prenatal testing. *Proc. Natl. Acad. Sci.* 115, E5106–E5114. doi: 10.1073/pnas.1804134115
- Taly, V., Pekin, D., Benhaim, L., Kotsopoulos, S. K., Le Corre, D., Li, X., et al. (2013). Multiplex picodroplet digital PCR to detect KRAS mutations in circulating DNA from the plasma of colorectal cancer patients. *Clin. Chem.* 59, 1722–1731. doi: 10.1373/clinchem.2013.206359
- Thierry, A. R., Mouliere, F., Gongora, C., Ollier, J., Robert, B., Ychou, M., et al. (2010). Origin and quantification of circulating DNA in mice with human colorectal cancer xenografts. *Nucleic Acids Res.* 38, 6159–6175. doi: 10.1093/nar/gkq421
- Thierry, A. R., El Messaoudi, S., Gahan, P. B., Anker, P., and Stroun, M. (2016). Origins, structures, and functions of circulating DNA in oncology. *Cancer Metastasis Rev.* 35, 347–376. doi: 10.1007/s10555-016-9629-x
- Tug, S., Helmig, S., Menke, J., Zahn, D., Kubiak, T., Schwarting, A., et al. (2014). Correlation between cell free DNA levels and medical evaluation of disease progression in systemic lupus erythematosus patients. *Cell. Immunol.* 292, 32–39. doi: 10.1016/j.cellimm.2014.08.002
- Umetani, N., Giuliano, A. E., Hiramatsu, S. H., Amersi, F., Nakagawa, T., Martino, S., et al. (2006). Prediction of breast tumor progression by integrity of free circulating DNA in serum. *J. Clin. Oncol.* 24, 4270–4276. doi: 10.1200/JCO.2006.05.9493
- Underhill, H. R., Kitzman, J. O., Hellwig, S., Welker, N. C., Daza, R., Baker, D. N., et al. (2016). Fragment length of circulating tumor DNA. *PLoS Genet.* 12, e1006162. doi: 10.1371/journal.pgen.1006162
- US Food & Drug Administration. (2016). Premarket approval P150044 — Cobas EGFR MUTATION TEST V2. FDA <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P150044>.
- Wan, J. C. M., Massie, C., Garcia-Corbacho, J., Mouliere, F., Brenton, J. D., Caldas, C., et al. (2017). Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nat. Rev. Cancer* 17, 223–238. doi: 10.1038/nrc.2017.7
- Yu, S. C. Y., Chan, K. C. A., Zheng, Y. W. L., Jiang, P., Liao, G. J. W., Sun, H., et al. (2014). Size-based molecular diagnostics using plasma DNA for noninvasive prenatal testing. *Proc. Natl. Acad. Sci.* 111, 8583–8588. doi: 10.1073/pnas.1406103111

Conflict of Interest: Authors JL, WTW, YH, CQ, LS, SY, KZ, GZ, HP, QL, YP, WWW, DY, SX and GT were employed by the company Geneis (Beijing) Co. Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Liu, Lang, Li, Wang, Peng, Wang, Han, Qi, Song, Yang, Zhang, Zang, Pei, Lu, Peng, Xi, Wang, Yuan, Bing, Zhou and Tian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.