EBioMedicine 55 (2020) 102462

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.elsevier.com/locate/ebiom

Research paper

Circulating tumour-derived *KRAS* mutations in pancreatic cancer cases are predominantly carried by very short fragments of cell-free DNA

Maria Zvereva[a,b], Gabriel Roberti[a,c], Geoffroy Durand[a], Catherine Voegele[a], Minh Dao Nguyen[a], Tiffany M. Delhomme[a], Priscilia Chopard[a], Eleonora Fabianova[d], Zora Adamcakova[d], Ivana Holcatova[e], Lenka Foretova[f], Vladimir Janout[g], Paul Brennan[a], Matthieu Foll[a], Graham B. Byrnes[a], James D. McKay[a], Ghislaine Scelo[a], Florence Le Calvez-Kelm[a,*]

[a] *International Agency for Research on Cancer (IARC), Genetic Cancer Susceptibility group, 150 Cours Albert Thomas, 69372 Lyon, France*
[b] *Faculty of Chemistry, Lomonosov Moscow State University, Moscow, Russian Federation*
[c] *Santa Casa de Sao Paulo of medical Sciences, Sao Paulo, Brazil*
[d] *Regional Authority of Public Health, Banska Bystrica, and Faculty of Health, Catholic University, Ružomberok, Slovakia*
[e] *First Faculty of Medicine, Charles University of Prague, Institute of Hygiene and Epidemiology, Prague, Czechia*
[f] *Masaryk Memorial Cancer Institute and Medical Faculty of Masaryk University, Brno, Czechia*
[g] *Faculty of Health Sciences, Palacky University, Olomouc, Czechia*

ARTICLE INFO

Article History:
Received 19 July 2019
Revised 13 September 2019
Accepted 22 September 2019
Available online 3 April 2020

Keywords:
Cell-free DNA
KRAS mutations
Plasma
Pancreatic cancer detection

ABSTRACT

Background: The DNA released into the bloodstream by malignant tumours· called circulating tumour DNA (ctDNA), is often a small fraction of total cell-free DNA shed predominantly by hematopoietic cells and is therefore challenging to detect. Understanding the biological properties of ctDNA is key to the investigation of its clinical relevance as a non-invasive marker for cancer detection and monitoring.
Methods: We selected 40 plasma DNA samples of pancreatic cancer cases previously reported to carry a *KRAS* mutation at the 'hotspot' codon 12 and re-screened the cell-free DNA using a 4-size amplicons strategy (57 bp, 79 bp, 167 bp and 218 bp) combined with ultra-deep sequencing in order to investigate whether amplicon lengths could impact on the capacity of detection of ctDNA, which in turn could provide inference of ctDNA and non-malignant cell-free DNA size distribution.
Findings: Higher *KRAS* amplicon size (167 bp and 218 bp) was associated with lower detectable cell-free DNA mutant allelic fractions ($p < 0{\cdot}0001$), with up to 4·6-fold (95% CI: 2·6–8·1) difference on average when comparing the 218bp- and the 57bp-amplicons. The proportion of cases with detectable *KRAS* mutations was also hampered with increased amplicon lengths, with only half of the cases having detectable ctDNA using the 218 bp assay relative to those detected with amplicons less than 80 bp.
Interpretation: Tumour-derived mutations are carried by shorter cell-free DNA fragments than fragments of wild-type allele. Targeting short amplicons increases the sensitivity of cell-free DNA assays for pancreatic cancer and should be taken into account for optimized assay design and for evaluating their clinical performance.
Funding: IARC; MH CZ – DRO; MH SK; exchange program between IARC and Sao Paulo medical Sciences; French Cancer League.

1. Introduction

One of the most promising areas in translational cancer research is the potential clinical use of "liquid biopsy" as a reservoir of biomarkers for cancer detection and monitoring [1]. Tumour-derived mutations in circulating cell-free DNA fragments (cfDNA), called circulating tumour DNA (ctDNA), have been identified in body fluids of cancer patients at late and early stages but considerable challenges remain before it could be used as clinical markers [2]. The technical challenge is that ctDNA often represents a very small fraction of the total cfDNA, the majority of which been suggested to originate from hematopoietic cell death [3]. The development of next-generation sequencing (NGS) and other genomic technologies have partially overcome these technical limitations providing detection thresholds of mutant allelic fraction (MAF) of ctDNA as low as 0·01%. Nevertheless, research is required to better understand the biology of both ctDNA and cfDNA to ease the detection of the ctDNA fraction. Current

* Corresponding author.
E-mail address: lecalvezf@iarc.fr (F. Le Calvez-Kelm).

https://doi.org/10.1016/j.ebiom.2019.09.042

Research in context

Evidence before this study

Circulating tumour DNA (ctDNA) carries the genetic mutations of the tumour cells they originate from. However, ctDNA fragments are a small fraction of the total circulating cell-free DNA fragments (cfDNA), making the mutations difficult to detect and challenging the use of ctDNA as clinical markers. Current knowledge of the biological properties and mechanisms of release of cfDNA and ctDNA remains limited. Previous studies reported a peak at 166 base pairs (bp) for cfDNA fragments, with some evidence that ctDNA fragments are on average shorter than fragments of non-tumour origin. This would further complicate the detection of tumour-specific mutations in cfDNA, as shorter fragments may be more challenging to amplify and sequence. In pancreatic cancer patients, a broad range of cfDNA fragment size was reported. Few months ago, Liu et al. showed that an unconventional single-strand library preparation associated with deep sequencing enhanced the detection of the *KRAS* oncogenic hotspot mutation in plasma of pancreatic patients.

Added value of this study

Here we tested 40 plasma samples of pancreatic cancer cases previously reported to carry *KRAS* mutation at the 'hotspot' codon 12. We use a four-size amplicon strategy (57 bp, 79 bp, 167 bp, and 218 bp) to re-screen the cfDNA with ultra-deep sequencing. We showed that the shorter the amplicon size, the higher the mutant allelic fraction. As a result, the proportion of cases with detectable *KRAS* mutations was multiplied by two for amplicons shorter than 80pb as compared to 218pb.

Implications of all the available evidence

Our study brings further evidence that tumour-derived mutations are predominantly carried by short and ultra-short cfDNA fragments. Hence, the specific biological fragmentation characteristics of ctDNA should be taken into account when performing liquid biopsies based assays. It is expected that enrichment of short cfDNA mole- cules would increase the ability to detect tumour-derived mutations in liquid biospsies, thus enhancing the clinical performance of the assays.

knowledge of the biological properties of cfDNA and ctDNA and the mechanisms of their release remains limited. While apoptosis and necrosis of cells have been implicated as possible sources of cfDNA in body fluids [4], recent data from cultured cell lines indicate that cfDNA could be actively secreted [5]. In the bloodstream, cfDNA exists as either free or associated with extracellular vesicles such as exosomes and these structures may also contribute to the mechanisms of release of DNA [6]. The size distribution of ctDNA and cfDNA may therefore reflect their heterogeneous origins and related cellular mechanisms involved in their release.

The detection of tumour-derived mutations in plasma samples with qPCR assays gave the first indications of the potential existence of shorter ctDNA than normal cfDNA [7,8]. Massively parallel sequencing brought evidence that the distribution of cfDNA fragment lengths displayed a most prominent peak at 166 bp in hepatocellular carcinoma patients, healthy subjects and HBV-carriers. However, cfDNA in HCC patients with high tumour derived MAF appeared more fragmented than in healthy individuals suggesting that ctDNA is shorter than non-malignant cfDNA [9]. More recently, a study using human xenograft models and human melanoma and lung cancer plasma samples consolidated these findings showing that mutant allele (ctDNA) occurred more commonly at a shorter fragment length (134−144 bp) than the fragment length of wild-type allele (165 bp) and that ctDNA was associated with a strong 10 bp periodicity characteristic of nuclease-cleaved nucleosome activity [10]. The authors also showed that the size-selection for short cell-free DNA fragment increased the mutant allele fraction of the EGFR T790M in human lung cancer cfDNA samples, suggesting that the experimental isolation of a specific subset of fragment lengths may improve sensitivity of detection of ctDNA. It has been also reported that in addition to short cfDNA fragments, blood from cancer patients may also contain large cfDNA fragments (>10 000 bp) that may originate from catastrophic events such as necrosis [4]. Whether the size distribution of cfDNA is homogenous among cancer patients or varies according to tissue of origin or other physio-pathological states is unknown. However, there are indications that cfDNA of cancer patients harbour particular nucleosome footprints that could infer the tumour's tissue of origin [3].

In patients with pancreatic cancer, various length measurements of plasma cfDNA have been specifically observed [11]. In a recent study, Liu and colleagues showed using a single-strand DNA library preparation and deep sequencing that the ctDNA fragmentation pattern was different between patients with advanced and early pancreatic cancers, the latter displaying more shortened ctDNA fragments carrying the *KRAS* oncogenic highly prevalent mutations [12].

We recently conducted the largest case-control study investigating the utility of *KRAS* mutations in plasma cfDNA samples as non-invasive biomarkers for the detection of pancreatic cancer, using a short amplicon-based (79 bp) deep sequencing strategy combined with analytical pipeline specifically designed for the detection of low-abundance mutations [13]. We selected *KRAS* codon 12 mutated cases to assess whether MAFs detected by amplicon-based deep sequencing depend on amplicon length due to differences in biological release or processing of normal cfDNA and ctDNA in pancreatic cancer, which in turn may affect the performance of ctDNA detection as suggested by Liu et al. [12].

2. Materials and methods

2.1. Study population and sample selection

In a multi-centre case-control study [14,15], we previously applied a *KRAS* amplicon-based deep sequencing strategy and an analytical pipeline designed for the detection of low-abundance mutations to screen plasma samples of 437 pancreatic cancer cases, 141 chronic pancreatitis subjects, and 394 healthy controls. We detected mutations in 21·1% ($N = 92$) of cases, of whom 82 (89·1%) carried at least one mutation at hotspot codons 12, 13 or 61 [13]. The remaining cfDNA samples were kept frozen at −20 °C for 2 years. We selected pancreatic cancer cases positive for *KRAS* codon 12 mutation in their plasma samples and for which sufficient amount of cfDNA was available after the first deep-sequencing screening to perform the ultra-deep sequencing of 4 independent amplicons ($N = 40$ out of the 74 pancreatic cancer patients in the original study). The proportion of cases carrying a codon 12 mutation stratified by recruitment centre, tumour histology types, stages and mutation types was not different between the initial and current study (appendix p1).

2.2. Sample processing and cfDNA KRAS mutation detection by ultra-deep sequencing

Collection and processing of peripheral blood from patients as well as cfDNA isolation and quantification are described in the initial case-control study [13]. Primers in the original study were designed to amplify partial KRAS exon 2 spanning codon 12 so that the amplicon size is 79 bp. For this study, we designed three additional amplicons of 57 bp, 167 bp and 218 bp so that shorter sizes overlap the sequence of amplicons of larger sizes (appendix p2). Fragment sizes

were defined so that it allows the intra- and inter-subject mutational load comparisons of smaller (57 bp, 79 bp) and larger (218 bp) fragment sizes to the averaged reported normal cfDNA size (167 bp).

Independent PCR amplifications using same experimental conditions were performed using 2 ng of cfDNA and barcoded libraries were prepared as previously described [13]. Emulsion amplification was performed on the Ion OneTouch 2 system using 7 μL of 100 pM library and the Ion PI Hi-Q OT2 200 Kit (Thermoscience Fisher). The sequencing reaction was performed on an Ion Proton System (Thermoscience Fisher) using Life Technologies' Ion PI™ Chip Kit v3 and Ion PI™ Hi-Q™ Sequencing 200 Kit (Thermoscience Fisher), according to the manufacturer's instructions.

2.3. Bioinformatics and statistical analyses

BAM files were generated by the Ion Torrent Proton server using default parameters (the Torrent Suite v5.8.0) and reads were mapped to the human hg19 genome using targeting sequencing default parameters for Ion PI™ Chip (ion_P10.1.17_tagSeq). Reads with mapping and base quality <20 were excluded from subsequent analyses.

We analysed the association between allelic fraction of *KRAS* mutant and length of the amplicon by applying a mixed-effects negative binomial regression, with the number of mutated reads for each amplicon as the response variable and the total number of corresponding reads as the offset variable. Number of reads were included in the models as integer (count) variables. To take into account the repeated measures per subject, e.g. the measures obtained for each of the four evaluated amplicon size, we used a random intercept for each subject, with the total number of reads used as offset. We provided rate ratio estimates and their 95% confidence intervals for the larger three amplicon sizes using the smallest amplicons as the reference. Overall p-values for the effect of amplicon size were calculated using the Wald test. Results were stratified by levels of allelic fractions detected using the 57 bp amplicon as the reference.

We used Needlestack, a variant caller algorithm suitable for the detection and quantification of low-abundance mutations [13,16,17] in multiple samples. In addition to the sample set, we sequenced 20 cfDNA samples from controls and selected to be negative after Needlestack analysis in the initial study to provide sufficient power to the negative binomial regression. We calculated for each sample a p-value for being a variant (outlier from the regression) that we further transformed into q-values to account for multiple testing. q-values are reported in Phred scale $Q = -10 \log_{10}$ (QVAL), and we used a threshold of $Q > 30$ to call variants, such as applied in the original study.

The Fisher exact test was used to evaluate whether the KRAS amplicon size was correlated with the ability to detect *KRAS* mutation in plasma DNA of pancreatic cancer cases.

3. Results

Forty pancreatic cancer cases positive for *KRAS* codon 12 mutations in their plasma DNA following Ion Torrent PGM 79-bp amplicon based sequencing [13] were selected to investigate whether the amplicon size may influence the ability to detect cfDNA *KRAS* mutations, which in turn may inform on the relative proportion of ctDNA and normal cfDNA in blood of cancer patients. The study was designed with the intent to explore whether smaller size amplicons, which could amplify any cfDNA fragments of the size or longer than the amplicons tested, would generate higher or lower *KRAS* mutant allelic fraction readout than the longer ones. Four amplicons were amplified from the same plasma DNA sample for ultra-deep sequencing and Needlestack variant calling (c.34 *G*>*C*, p.G12R; c.35 *G*>*T*, p.G12V; c.35 *G*>*A*, p.G12D). The sizes of amplicons 57 bp, 79 bp, 167 bp and 218 bp were selected for comparison of short, ultra-short DNA fragments lengths to the reported the most prominent peak size of cell-free DNA (166 bp) and larger fragments.

Needlestack analysis was conducted independently for each amplicon and for each nucleotide change. Methodologically, we previously showed that our amplicon-based deep sequencing approach of minimal amount of plasma cfDNA (2 ng) combined the Needlestack 210 analysis could detect *KRAS* p.G12V mutation down to a minor allele frequency of 0·2% at a Phred scale q-value above 30 (QVAL>30) and a read depth of approximately 2500 reads [13]. It therefore represents a reliable alternative to methods established to minimize sequencing errors such as attaching unique molecule identifiers (UMI) to DNA. In this study, similarly samples with QVAL>30 were considered as outliers of the regression and therefore mutated (appendix p3). We first verified that the detection capacity of *KRAS* mutations was not altered by the between-amplicon variability of sequencing error rates for the three considered mutation types (appendix p3) or of sequencing read depth (appendix p4). Robust estimation of averaged sequencing error rates for each amplicon and for each nucleotide change were found to be low (in the range of 10^{-5} for c.34G>*C*, and 10^{-4} c.35G>*A* and c.35G>*T* error rates) and consistent across amplicons for the same mutation (appendix p3), indicating that various DNA fragments length sequenced have a minimal impact on technical sequencing variability. The average read depth and standard deviation were 48 392 (2043), 39 062 (12 836), 35 097 (14 592), and 21 507 (12 222) for 57 bp, 79 bp, 167 bp and 218 bp amplicon sizes, respectively. To check for the influence of sequencing depth on the robustness of *KRAS* mutation detection threshold, our set of sequencing data was used to predict the theoretical lowest AFs that could be detected for the three mutation types and four evaluated amplicons according to increasing sequencing depth and different fixed QVALs of 10, 30, 50, 70 and 100 (appendix p4). Within-amplicon prediction plots are almost stackable for the four amplicons and the minimal detectable AFs are minimally influenced by sequencing depth higher than 15 000X. Considering the high average sequencing read depth obtained for the evaluated plasma DNA series, the minimal within-sample variability in sequencing depth should not hamper the robustness of low AF mutation detection. Table 1 provides the list of mutations initially reported and their detected MAF for various sizes of KRAS amplicons in the current experiments. To investigate whether there were significant differences in allelic fractions of KRAS mutated DNA for different lengths of amplicons as suggested by Fig. 1, we used a mixed negative binomial regression model on all data of the 40 cases (overall read depth and alternative mutated reads). Rate ratios and 95% confidence intervals were calculated for the 57 bp, 79 bp, and 167 bp amplicons with the largest, 218bp-amplicon used as reference (Table 2). Compared with the 218bp-amplicon, detected MAFs were 1·8-fold higher (95%CI: 1·0–3·0) with the 167bp-amplicon, 4·1-fold higher (95%CI: 2·3–7·3) with the 79bp-amplicon, and 4·6-fold higher (95%CI: 2·6–8·1) with the 57-bp amplicon, indicating that *KRAS* mutations are more likely to be carried by shorter cfDNA fragments than the fragments of wild-type allele. However, there was no statistical difference between the two smallest amplicons ($p = 0{\cdot}67$). The rate ratio estimates tended to increase slightly when focusing the analysis on MAFs<20%, with a 6·1-fold difference (95% CI: 3·1–11·9) but remained stable when restricting the analysis to MAFs<10% and <5% (Table 2). This suggests that pancreatic cancer patients with MAF *KRAS* mutated DNA in their plasma lower than 20% carry more fragmented ctDNA relative to wild-type cfDNA, compared to those high fractions of tumour DNA. When stratifying by tumour stages, we observed similar increases in *KRAS* MAFs with smaller amplicon sizes as compared to non-stratified cases, suggesting that ctDNA of regional and systemic pancreatic cancer cases may undergo similar mechanisms of cleavage. MAFs obtained from Ion Torrent Proton data for the 79 bp amplicon and MAFs from the previous study using the same amplicon were consistent (r2 = 0·9783) (appendix pp.5–6). Applying a threshold of QVAL>30, *KRAS* mutations were identified in 35 of 40 cases in at least one of the four KRAS amplicons tested

Table 1
KRAS Mutant Allelic fractions according to amplicon size in plasma cfDNA of pancreatic cancer cases.

Sample ID	*KRAS* mutation	Amplicon size			
		57bp MAF (%)	79bp MAF (%)	167bp MAF (%)	218bp MAF (%)
CA80	c.35G>*A*; p.G12D	66·66	75·53	68·49	61·38
CA79	c.34G>*C*; p.G12R	58·32	65·76	27·12	70·78
CA73	c.35G>*T*; p.G12V	44·95	28·57	13·79	2·51
CA77	c.35G>*T*; p.G12V	38·36	27·28	21·39	30·69
CA75	c.34G>*C*; p.G12R	30·49	–	10·51	7·27
CA93	c.35G>*T*; p.G12V	30·46	28·09	25·07	15·64
CA74	c.35G>*A*; p.G12D	16·26	18·87	12·16	11·35
CA66	c.35G>*T*; p.G12V	13·58	11·52	***0·01***	***0·00***
CA68	c.35G>*T*; p.G12V	12·98	8·86	3·65	2·43
CA70	c.35G>*A*; p.G12D	12·19	13·39	5·25	3·35
CA49	c.35G>*A*; p.G12D	11·39	2·35	0·47	0·22
CA71	c.34G>*C*; p.G12R	10·40	11·69	12·03	7·00
CA50	c.35G>*T*; p.G12V	7·36	4·74	1·92	0·64
CA72	c.35G>*A*; p.G12D	7·26	10·66	4·03	***0·03***
CA57	c.35G>*T*; p.G12V	5·76	2·77	1·65	2·30
CA55	c.35G>*T*; p.G12V	3·86	4·28	6·22	2·17
CA61	c.35G>*A*; p.G12D	3·58	6·50	1·49	–
CA45	c.35G>*T*; p.G12V	2·26	2·81	2·41	1·78
CA43	c.35G>*T*; p.G12V	1·90	2·21	0·40	***0·04***
CA24	c.35G>*T*; p.G12V	1·83	0·39	1·29	0·36
CA44	c.35G>*A*; p.G12D	1·73	2·33	4·49	6·74
CA46	c.35G>*T*; p.G12V	1·50	2·13	0·47	–
CA58	c.35G>*A*; p.G12D	1·15	1·36	1·16	***0·01***
CA17	c.35G>*T*; p.G12V	1·13	1·30	***0·02***	***0·03***
CA28	c.35G>*T*; p.G12V	0·95	0·44	***0·02***	***0·00***
CA39	c.35G>*A*; p.G12D	0·82	2·94	0·54	0·59
CA19	c.35G>*T*; p.G12V	0·60	0·19	***0·00***	***0·01***
CA26	c.34G>*C*; p.G12R	0·54	0·48	***0·00***	***0·00***
CA31	c.35G>*T*; p.G12V	0·43	***0·03***	***0·01***	***0·03***
CA38	c.35G>*A*; p.G12D	0·41	2·25	0·59	0·61
CA25	c.35G>*A*; p.G12D	0·34	0·87	0·24	***0·01***
CA34	c.35G>*A*; p.G12D	0·24	1·05	0·45	–
CA22	c.34G>*C*; p.G12R	0·20	***0·01***	***0·01***	***0·00***
CA29	c.34G>*C*; p.G12R	***0·01***	3·78	***0·00***	***0·00***
CA16	c.35G>*T*; p.G12V	***0·02***	0·91	4·91	***0·15***
CA27	c.35G>*A*; p.G12D	***0·06***	***0·05***	***0·07***	***0·11***
CA11	c.34G>*C*; p.G12R	***0·00***	***0·01***	***0·01***	***0·01***
CA21	c.35G>*T*; p.G12V	***0·02***	***0·02***	***0·04***	***0·02***
CA15	c.34G>*C*; p.G12R	***0·00***	***0·00***	***0·00***	***0·00***
CA13	c.34G>*C*; p.G12R	***0·00***	***0·00***	***0·00***	***0·00***

No data: Bold Italic: *KRAS* mutation not detected, e.g. samples not outliers of the Needlestack regression at QVAL<30.

(Table 1). The low AF *KRAS* mutations (between 0·2% and 0·4%) previously detected in five samples (CA27, CA11, CA21, CA15, CA13) in the original study where the read depth was lower were not detected at QVAL>30.

The sequencing data of the 35 cfDNA *KRAS* mutation positive samples were used to estimate the capacity of targeted NGS-based detection of cfDNA *KRAS* mutations according to amplicon size (Fig. 2). The proportion of detected pancreatic cancer cases using the 57 bp or 79 bp *KRAS* amplicon in cfDNA plasma samples was comparable (94·2% and (94·1% respectively) but statistically higher than the proportion of subjects detected with the longer amplicons of 218 bp (59·6%; $p < 0{\cdot}001$), indicating that targeting short amplicon is key to increase the sensitivity of cfDNA assays.

Finally, we classified samples based on their MAF patterns using the level of difference in MAF between 57 bp and 218 bp amplicons in samples carrying a *KRAS* mutation in at least one amplicon ($N = 32$). The most prevalent pattern characterized by a minimum decreased MAFs of 40% according to *KRAS* amplicon length was observed in 21 cases of 32 (65·6%) (Fig. 3) ranging from 43·9% (CA55) to more than 99·9% (CA66) indicating that shorter forms of tumour-derived cfDNA (ctDNA) are likely to be the results of specific nucleosome-wrapping and cleavage mechanisms that are different to the mechanisms producing cfDNA originating from hematopoietic cell death. An opposite pattern as compared to the most prevalent one, e.g. displaying a progressive increase in MAF with amplicon length was observed in one patient (CA44), suggesting that, in some rare cases of pancreatic cancer with particular physio-pathological states, wild-type fragments released in the blood stream (from the tumour or from the micro-environment of the tumour) may be more prone to fragmentation than mutated ctDNA. We can not exclude that longer mutated tumour fragments may be preferentially released in the blood circulation but our assay was not designed to assess this parameter, as amplicons of small sizes should amplify all cfDNA fragments of equal or larger sizes than the amplicons. Finally, while there was no statistical difference overall between *KRAS* mutant allelic fractions of the 57-bp and 79-bp amplicons, we observed 11 samples (CA73, CA77, CA68, CA49, CA50, CA57, CA24, CA28, CA19, CA31 and CA22) with substantial decreased MAF (Table 1) ranging from 28·9% to 95·0%, demonstrating that ultra-short ctDNA of fragment sizes, between 57 bp and 79 bp, do exist in some pancreatic cancer patients.

4. Discussion

Our study was designed to assess the impact of four amplicon sizes on the *KRAS* codon 12 mutation detection using 40 reported mutated plasma samples of pancreatic cancer cases [13] and ultra-deep sequencing. We aimed to explore whether smaller size amplicons (<80 bp), which are generally below what is generally used by 316 many NGS-based cfDNA assays, would generate higher *KRAS* mutant allelic fraction readout than the longer ones, with the intent to provide inference of the cfDNA and ctDNA size profiles and of sensitivity of ctDNA assays. We showed that the ability to detect *KRAS* mutations in positive cfDNA samples is hampered with increased sizes of *KRAS* amplicons (167 bp and 218 bp), the latter being significantly correlated with substantial decrease of MAFs. Altogether, our results highlight that ctDNA originating from pancreatic tumours and carrying tumour-derived alterations are significantly shorter than non-malignant cfDNA fragments, supporting previous studies reporting similar findings in human plasma samples of melanoma, lung cancer and hepatocellular carcinoma patients and in human cancer xenografts models [7–10,18] and more recently in plasma samples of pancreatic cancer patients [12].

Our observation that targeting shorter cfDNA fragments (<167 bp) with the 57 bp and 79 bp assay substantially increased the cancer-associated MAF is in agreement with previous work reporting that ctDNA from xenograft model display a principal fragment length at 134–144 bp and an animal cell-free DNA at 167 bp [10]. We also report the existence of ctDNA fragments sizes between 57 bp and 79 bp in some cases with pancreatic regional or systemic tumour stages. Interestingly, Liu and colleagues recently reported that shortened ctDNA fragments of this size range were more preferably found in patients with early-stage pancreatic tumours rather than in patients with more advanced stages [12]. Such understudied forms of ultra-short human cell-free DNA originating from mitochondrial and short nuclear genomic cfDNA have been also recently uncovered in plasma of lung transplants recipients (<100 bp) [5]. Our results also corroborate that it exists a specific ctDNA cleavage mechanism, which is different to the shortening process applied to normal cell-free DNA originating from hematopoietic cells and that seems to reflect the nuclease-cleaved nucleosome activity [10]. Snyder et al. described that the global nucleosome-associated DNA packing of cell-free DNA so called 'nucleosome footprint' could infer cell types contributing to cfDNA in cancer patients, suggesting that cleavage-associated patterns are specific to the tissue where the cancer originated [3]. However, growing evidence supported by our results and previous reports demonstrate that ctDNA shortening happens in patients with many cancer types. To what extent the nucleosome-associated cleavage pattern of blood ctDNA differs according to the cancer tissues-of-origin and how it could influence mutation

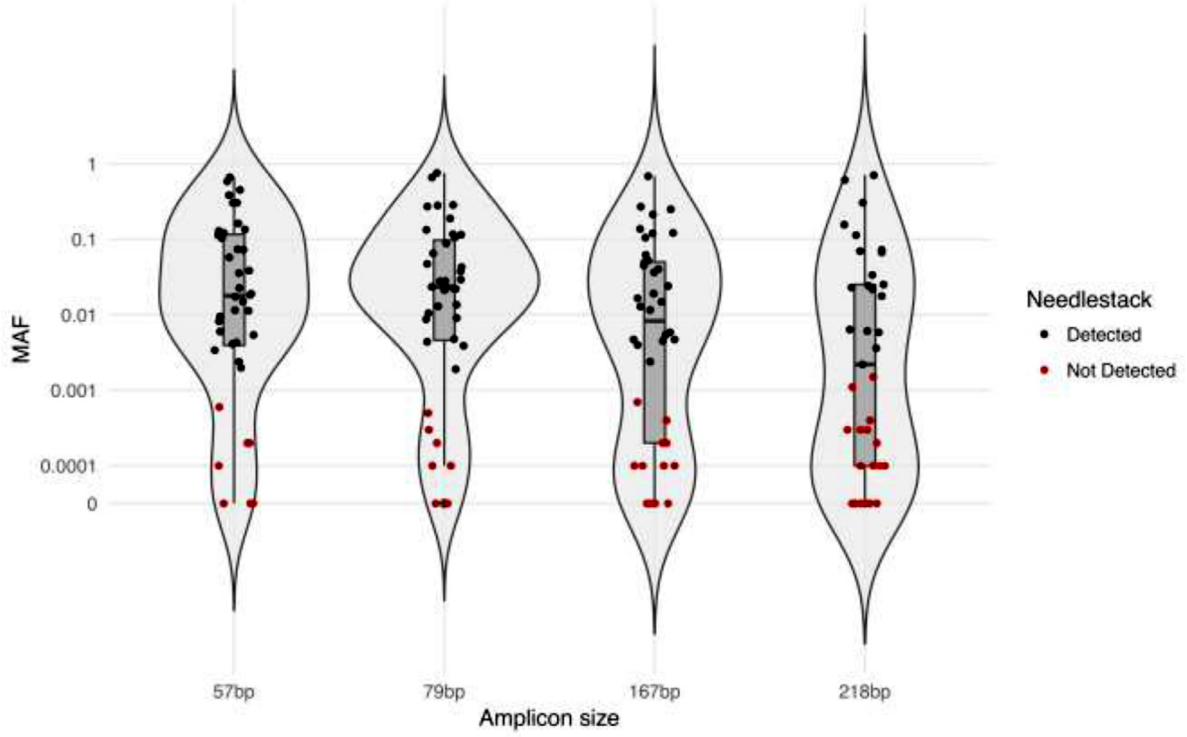

Fig. 1. Violin plot of *KRAS* MAF in pancreatic cancer cases according to amplicon lengths. Grey dots: MAFs of cases with detected mutations after Needlestack analysis (QVAL>30); red dots: Sequencing noise signal of cases with undetected mutations after Needlestack analysis (QVAL<30).

Table 2
Association between the amplicon length and the *KRAS* Mutant Allelic Fraction in plasma cfDNA of pancreatic cancer cases.

	KRAS amplicon size (bp)	Rate Ratio	RR 95% CI	*P*-value
Overall				
All (*N* = 40)	218	ref	–	
	167	1·8	1·0–3·0	
	79	4·1	2·3–7·3	<0·0001
	57	4·6	2·6–8·1	
KRAS mutant allelic fraction (MAF) for the 57 bp amplicon				
MAF<20% (*N* = 34)	218	ref	–	
	167	2·2	1·1–4·0	
	79	65·6	2·9–10·8	<0·0001
	57	6·1	3·1–11·9	
MAF<10% (*N* = 28)	218	ref	–	
	167	2·5	1·2–5·1	
	79	6·1	2·9–13·2	<0·0001
	57	6·0	2·8–12·8	
MAF<5% (*N* = 25)	218	ref	–	
	167	2·5	1·1–5·5	
	79	6·2	2·6–14·4	0·0003
	57	5·8	2·4–13·7	
Stages*				
Regional (*N* = 8)	218	ref	–	
	167	2·2	0·8–5·5	
	79	4·8	1·9–12·5	0·0018
	57	4·5	1·7–11·7	
Systemic (*N* = 21)	218	ref	–	
	167	1·5	0·7–3·5	
	79	4·2	1·7–10·4	0·0005
	57	4·3	1·8–10·6	
Unknown (*N* = 11)	218	ref	–	
	167	1·9	0·7–4·7	
	79	3·1	1·2–8·2	0·0060
	57	4·5	1·7–12.27	

* Stage grouping was defined as regional, and systemic cancers, based on TNM staging (AJCC 6th edition) when available, and estimation by the clinician when formal TNM staging was not available or not complete.

detection is not known. Similar cleavage patterns with a prominent peak at 143 bp have been observed for cfDNA of placental origin in contrast to the peak at 166 bp for the maternal cfDNA [19,20]. Similar observations have been made for transplantation [21] and immune-related disease [22], further supporting the idea that cleavage process of cfDNA originating from non-hematopoietic cells in general is different to that of hematopoietic cells.

Our observation that the majority of pancreatic cancer cases carry *KRAS* mutations in small fragments of cfDNA are indicative of a prominent role of apoptosis as mechanism of ctDNA release in the bloodstream and consistent with results of previous studies [9,10]. Further, the level of decreased MAF with increased amplicon length may reflect the level of contribution of the tumour apoptotic process to the release of short, fragmented ctDNA molecules. However, Jiang et al. reported using copy number analysis by paired-end sequencing that HCC patients with low fraction of tumour DNA in plasma had 373 an overrepresentation of larger-size fragments, which is difficult to reconcile with our findings where we observed an over representation of fragmented ctDNA in patients with low AF KRAS mutated DNA in their plasma as compared to patients with high fractions of total tumour DNA in their plasma samples. While our assay is not designed to evaluate the contribution of larger fragments to the ctDNA load, we could speculate that in a minor fraction of pancreatic cancer patients, the less pronounced decreased allelic fraction may reflect the almost equal release of short and long mutated and wild-type cfDNA fragments which could in turn may be the result of concomitant apoptotic and necrotic processes, such as already reported in pancreatic cancer patients [11]. More recently, double stranded genomic DNA spanning all chromosomes with mutated *TP53* and *KRAS* DNA in serum exosomes of pancreatic cancer patients [23] have been described and active secretion of tumour-derived exosomes into the bloodstream can be an additional explanation of the presence of long fragments in addition to the necrotic process. Our results indicate however that the generation of long ctDNA plasma molecules is a minority event as compared to DNA fragmented patterns.

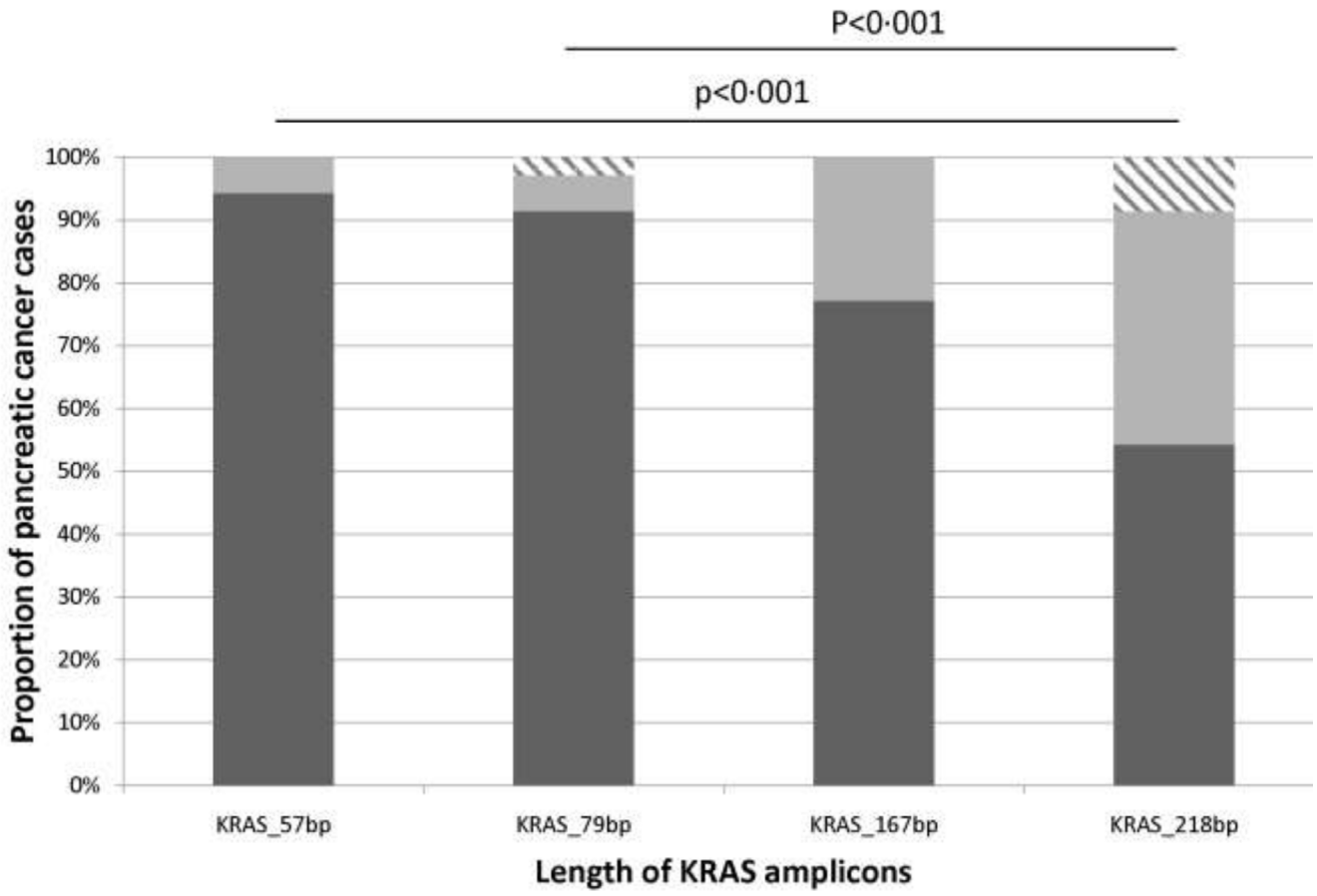

Fig. 2. Capacity of targeted ultra-deep sequencing detection of cfDNA *KRAS* mutations according to amplicon lengths in cell free DNA samples of pancreatic cancer cases. Dark Grey: cases with detected mutation; Light grey: Cases with undetected mutation; dotted line: Missing data. The proportions of detected mutant samples were compared between each amplicon size with the Fisher exact test. Two-tailed p-values<0·05 are shown.

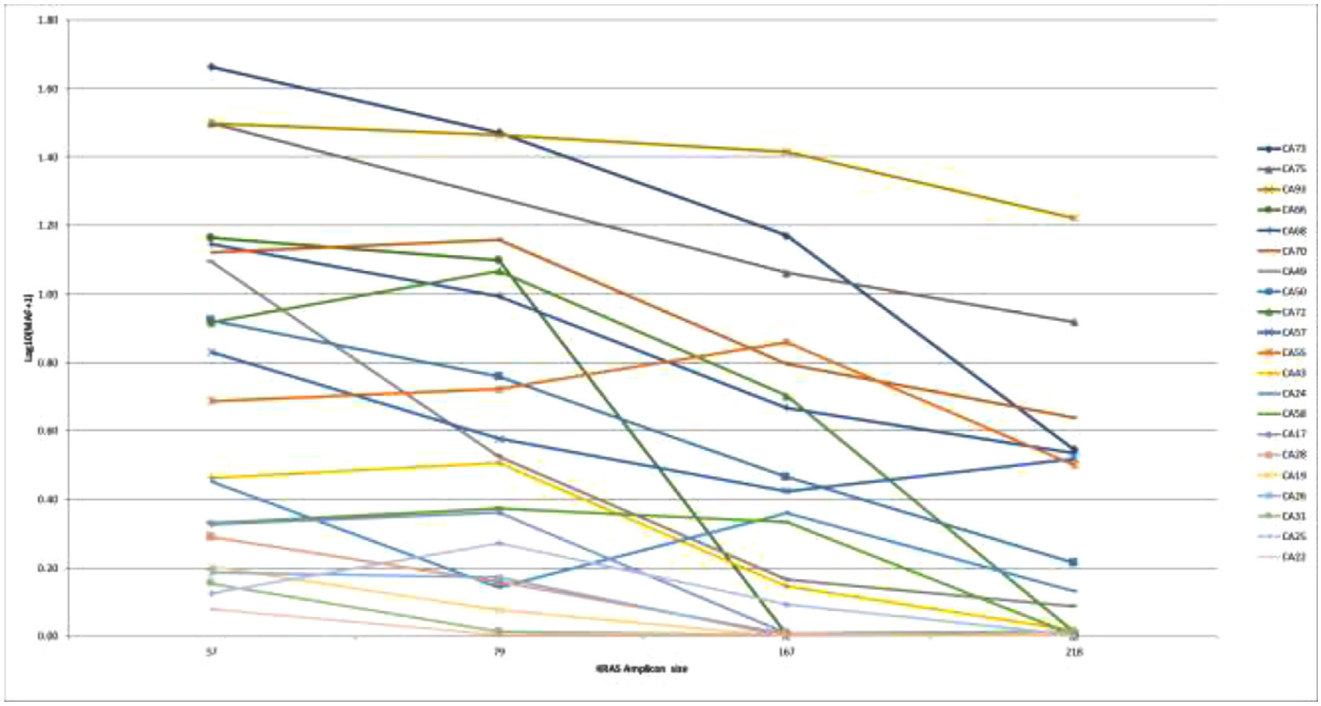

Fig. 3. Mutant Allelic Frequency pattern of the majority of pancreatic cancer cases according to amplicon lengths. Log (MAF+1) of samples carrying a *KRAS* mutation in at least one amplicon and with a level of decreased MAF between 57 bp and 218 bp amplicons > 40% ($N = 21$).

Regardless of the mechanisms involved in the generation of circulation DNA in the bloodstream, we proved that targeting short *KRAS* amplicons (57 bp or 79 bp) significantly improved NGS-based *KRAS* mutation detection by enriching the cfDNA MAF in pancreatic cancer patients, confirming similar results obtained with alternative approaches such as *KRAS* short amplicons (85 bp versus 120 bp) qPCR-based assay in a series of *KRAS* mutated colorectal patients [24] and single-strand library preparation combined with deep-sequencing in a series of pancreatic cancer patients [12]. Alternative methods for enrichment of small circulating DNA fragments (<150 bp) over cfDNA have been reported to improve sensitivity in detecting circulating genetic alterations. I*n vitro* size selection methods based on gel extraction platforms demonstrated improved detection of tumour-derived mutations of more than 2-fold [25,26] and magnetic beads-based methods showed improved detection of cfDNA of placental origin in prenatal non-invasive assay [27]. More recently, an innovative enrichment approach associating ion-tagged oligonucleotides to magnetic ionic solvents managed to pre-concentrate *KRAS* mutated fragments in plasma samples and could be applied towards ctDNA extraction [28]. Appropriate assay designs allowing for enrichment of short tumour-derived fragments over other longer non-malignant DNA fragments should be evaluated in large case-controls series to evaluate whether a significant increased clinical sensitivity in liquid biopsy-based cancer detection could be expected.

One limitation of our study is that we selected cfDNA *KRAS* mutated samples based on their positivity using the 79 bp assay to get insight into mechanisms of release of ctDNA into the blood stream. While there was no statistical difference in MAF of ctDNA between 57-bp and 79-bp assays overall in pre-selected 79-bp mutated cases, it would be interesting to screen negative cases for 79-bp assays with enough remaining cfDNA with the 57bp-assay to assess whether an improved clinical sensitivity of KRAS circulating mutations in detecting pancreatic cancer (>21%) could be reached. This was suggested by Liu et al. who reported a detection rate of 70% of 112 pancreatic cancer cases with the detection of KRAS mutation in plasma samples using single-strand library preparation combined with deep-sequencing [12]. Another limitation of our study is the 2-year storage of cfDNA samples at −20°C between the initial Ion Torrent PGM-based and second Proton-based screening and about 10 times increased sequencing coverage using the Ion Torrent Proton. One can not exclude that the absence of confirmation of the 5 low-level mutated cases samples originates from false positive calls in the initial screening or more likely from false negatives calls due to a preferential degradation of low-level ultra-short ctDNA fragments over time during storage or random sampling of low-level ctDNA with absence of mutated molecules in the proton screening. The latter two hypothesis are further supported in those cases by the confirmation of *KRAS* mutations in two additional cases at only 57 bp and by the fact that 11/16 cases with 79-bp MAF<1% in the initial study have lower MAFs or undetectable KRAS mutations using the 79bp-assay in the current study. However, overall there was a good concordance between the MAFs obtained from the 79 bp amplicons using the 2 NGS instruments (r^2 = 0·9783) and the degradation of ctDNA overtime was not seen in 18/23 cases with MAFs>1% (appendix pp. 5–6). Overall, those limitations do not alter the relevance of our results regarding the generation of ultra-short ctDNA molecules in the majority of pancreatic cancer cases for which the application of appropriate assay designs better discriminating ctDNA over non-malignant cfDNA in general should lead to higher diagnostic value of blood-based liquid biopsies.

Declaration of Competing Interest

We declare no competing interests.

Acknowledgments

The authors would like to thank all patients participating in the multi-centre case-control study and the clinical research team members who supported the sample collection. The authors thank the Ministry of Health of the Czech Republic MH CZ – DRO, the Ministry of Health of the Slovak Republic, the French Cancer League and the exchange program between IARC and Santa Casa de Sao Paulo of medical Sciences. The funders had no role in study design, data collection, data analysis, interpretation or writing of the report.

Funding sources

IARC; the recruitment of subjects was supported by the Ministry of Health of the Czech Republic MH CZ - DRO (MMCI, 00209805) and by the Ministry of Health of the Slovak Republic (Contract no. 2007/17-RUVZBB-02). GB was supported by an exchange program between IARC and Santa Casa de Sao Paulo of medical Sciences. MDN and TMD were supported by the French Cancer League.

Disclaimer

Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.ebiom.2019.09.042.

References

[1] Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A. Liquid biopsy: monitoring cancer-genetics in the blood. Nat Rev Clin Oncol 2013;10:472–84.
[2] Schwarzenbach H, Hoon DS, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. Nat Rev Cancer 2011;11:426–37.
[3] Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin. Cell 2016;164:57–68.
[4] Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. Cancer Res. 2001;61:1659–65.
[5] Bronkhorst AJ, Wentzel JF, Aucamp J, van Dyk E, du Plessis L, Pretorius PJ. Characterization of the cell-free DNA released by cultured cancer cells. Biochim. Biophys. Acta 2016;1863:157–65.
[6] Thierry AR, El Messaoudi S, Gahan PB, Anker P, Stroun M. Origins, structures, and functions of circulating DNA in oncology. Cancer Metastasis Rev 2016;35:347–76.
[7] Diehl F, Li M, Dressman D, He Y, Shen D, Szabo S, et al. Detection and quantification of mutations in the plasma of patients with colorectal tumors. In: Proceedings of the National Academy of Sciences of the United States of America, 102; 2005. p. 16368–73.
[8] Mouliere F, El Messaoudi S, Gongora C, Guedj AS, Robert B, Del Rio M, et al. Circulating cell-free DNA from colorectal cancer patients may reveal high KRAS or BRAF mutation load. Transl Oncol 2013;6:319–28.
[9] Jiang P, Chan CW, Chan KC, Cheng SH, Wong J, Wong VW, et al. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. In: Proceedings of the National Academy of Sciences of the United States of America, 112; 2015. p. E1317–25.
[10] Underhill HR, Kitzman JO, Hellwig S, Welker NC, Daza R, Baker DN, et al. Fragment length of circulating tumor DNA. PLoS Genet. 2016;12:e1006162.
[11] Giacona MB, Ruben GC, Iczkowski KA, Roos TB, Porter DM, Sorenson GD. Cell-free DNA in human blood plasma: length measurements in patients with pancreatic cancer and healthy controls. Pancreas 1998;17:89–97.
[12] Liu X, Liu L, Ji Y, Li C, Wei T, Yang X, et al. Enrichment of short mutant cell-free DNA fragments enhanced detection of pancreatic cancer. EBioMedicine 2019;41:345–56.
[13] Le Calvez-Kelm F, Foll M, Wozniak MB, Delhomme TM, Durand G, Chopard P, et al. KRAS mutations in blood circulating cell-free DNA: a pancreatic cancer case-control. Oncotarget 2016;7:78827–40.
[14] Brenner DR, Wozniak MB, Feyt C, Holcatova I, Janout V, Foretova L, et al. Physical activity and risk of pancreatic cancer in a central European multicenter case-control study. Cancer causes & control: CCC 2014;25:669–81.

[15] Urayama KY, Holcatova I, Janout V, Foretova L, Fabianova E, Adamcakova Z, et al. Body mass index and body size in early adulthood and risk of pancreatic cancer in a central European multicenter case-control study. Int J Cancer 2011;129:2875–84.
[16] Fernandez-Cuesta L, Perdomo S, Avogbe PH, Leblay N, Delhomme TM, Gaborieau V, et al. Identification of circulating tumor dna for the early detection of small-cell lung cancer. EBioMedicine 2016;10:117–23.
[17] Github IARCbioinfo/Needlestack: Multi-sample somatic variant caller.
[18] Mouliere F, Robert B, Arnau Peyrotte E, Del Rio M, Ychou M, Molina F, et al. High fragmentation characterizes tumour-derived circulating DNA. PLoS One 2011;6:e23418.
[19] Lo YM, Chan KC, Sun H, Chen EZ, Jiang P, Lun FM, et al. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. Sci Transl Med 2010;2:61ra91.
[20] Yu SC, Lee SW, Jiang P, Leung TY, Chan KC, Chiu RW, et al. High-resolution profiling of fetal DNA clearance from maternal plasma by massively parallel sequencing. Clin. Chem. 2013;59:1228–37.
[21] Zheng YW, Chan KC, Sun H, Jiang P, Su X, Chen EZ, et al. Nonhematopoietically derived DNA is shorter than hematopoietically derived DNA in plasma: a transplantation model. Clin. Chem 2012;58:549–58.
[22] Chan RW, Jiang P, Peng X, Tam LS, Liao GJ, Li EK, et al. Plasma DNA aberrations in systemic lupus erythematosus revealed by genomic and methylomic sequencing. In: Proceedings of the National Academy of Sciences of the United States of America, 111; 2014. p. E5302–11.
[23] Kahlert C, Fiala M, Musso G, Halama N, Keim S, Mazzone M, et al. Prognostic impact of a compartment-specific angiogenic marker profile in patients with pancreatic cancer. Oncotarget 2014;5:12978–89.
[24] Andersen RF, Spindler KL, Brandslund I, Jakobsen A, Pallisgaard N. Improved sensitivity of circulating tumor DNA measurement using short PCR amplicons. Clin. Chim. Acta 2015;439:97–101.
[25] Mouliere F, Chandrananda D, Piskorz AM, Moore EK, Morris J, Ahlborn LB, et al. Enhanced detection of circulating tumor DNA by fragment size analysis. Sci Transl Med 2018;10.
[26] Hellwig S, Nix DA, Gligorich KM, O'Shea JM, Thomas A, Fuertes CL, et al. Automated size selection for short cell-free DNA fragments enriches for circulating tumor DNA and improves error correction during next generation sequencing. PLoS One 2018;13:e0197333.
[27] Hu P, Liang D, Chen Y, Lin Y, Qiao F, Li H, et al. An enrichment method to increase cell-free fetal DNA fraction and significantly reduce false negatives and test failures for non-invasive prenatal screening: a feasibility study. J Transl Med 2019;17:124.
[28] Emaus MN, Clark KD, Hinners P, Anderson JL. Preconcentration of DNA using magnetic ionic liquids that are compatible with real-time PCR for rapid nucleic acid quantification. Anal Bioanal Chem 2018;410:4135–44.