Carrier Molecules and Extraction of Circulating Tumor DNA for Next Generation Sequencing in Colorectal Cancer

Martin Beránek1,2,*; Igor Sirák1; Milan Vošmík1; Jiří Petera1; Monika Drastiková1; Vladimír Palička1

1 Institute of Clinical Biochemistry and Diagnostics, University Hospital Hradec Králové, Charles University, Faculty of Medicine in Hradec Králové, Czech Republic
2 Department of Biochemical Sciences, Charles University, Faculty of Pharmacy in Hradec Králové, Czech Republic
3 Department of Oncology, University Hospital Hradec Králové, Charles University, Faculty of Medicine in Hradec Králové, Czech Republic
* Corresponding author: Institute of Clinical Biochemistry and Diagnostics, University Hospital Hradec Králové, Sokolská 581, 500 05 Hradec Králové, Czech Republic; e-mail: beranek@lfhk.cuni.cz

Summary: The aims of the study were: i) to compare circulating tumor DNA (ctDNA) yields obtained by different manual extraction procedures, ii) to evaluate the addition of various carrier molecules into the plasma to improve ctDNA extraction recovery, and iii) to use next generation sequencing (NGS) technology to analyze KRAS, BRAF, and NRAS somatic mutations in ctDNA from patients with metastatic colorectal cancer. Venous blood was obtained from patients who suffered from metastatic colorectal carcinoma. For plasma ctDNA extraction, the following carriers were tested: carrier RNA, polyadenylic acid, glycogen, linear acrylamide, yeast tRNA, salmon sperm DNA, and herring sperm DNA. Each extract was characterized by quantitative real-time PCR and next generation sequencing. The addition of polyadenylic acid had a significant positive effect on the amount of ctDNA eluted. The sequencing data revealed five cases of ctDNA mutated in KRAS and one patient with a BRAF mutation. An agreement of 86% was found between tumor tissues and ctDNA. Testing somatic mutations in ctDNA seems to be a promising tool to monitor dynamically changing genotypes of tumor cells circulating in the body. The optimized process of ctDNA extraction should help to obtain more reliable sequencing data in patients with metastatic colorectal cancer.

Keywords: Carrier; Extraction; Circulating tumor DNA; Next generation sequencing; Real-time PCR

Introduction

In mammalian cells, deoxyribonucleic acid saving genetic information is located in nucleus and mitochondria. Low amounts of genomic DNA are released into the blood plasma as cell-free DNA (cfDNA) with a median half-life of 16 minutes (1). In healthy subjects, the cfDNA concentration usually ranges between 0 ng/mL and 100 ng/mL. This corresponds to 0–15,150 genome equivalents per mL (GE/mL) (2). Numerous studies reported elevated cfDNA in pregnancy (fetal DNA), inflammation, autoimmune diseases, acute rejection of transplants, sepsis, or cancer (3–6), where circulating tumor-derived DNA (ctDNA) could reach hundreds of ng/mL (2, 7). In metastatic patients with increased ctDNA, the overall two-year survival rate of 48% was described (8).

Cell-free DNA is formed by 100–200 bp chromosomal fragments with the appropriate length of 311 nm (9). These short fragments were found in the plasma of both patients with malignancies or benign polyps, and/or healthy individuals (10, 11). Integral DNA molecules in plasma, on the other hand, originate from leukocytes or viable circulating tumor cells (12).

Previously published papers showed that somatic mutations in the KRAS (Kirsten rat sarcoma viral oncogene homolog) gene are often present in ctDNA of individuals suffering from pancreatic or gastrointestinal tumors (13). The mutations in KRAS codons 12, 13, and 61 were observed in plasma of one-fourth of metastatic cases, and an 80–86% mutation match between primary tumor tissues and ctDNA was demonstrated (14, 15).

Since the determination of the mutation status in the tumor tissue is necessary for the indication of targeted biological treatment of metastatic colorectal cancer, a panel of mutations tested in KRAS and other genes is being completed. In their analysis, sensitive investigation methods including real-time PCR, digital PCR, COLD PCR, reverse hybridization strips, or next generation sequencing have been applied. In this context, a clinical benefit of ctDNA testing is considered, as well. A reliable analytical process, however, requires relatively high volumes of plasma and the highest ctDNA concentrations in extracts possible.

The aims of the study were: i) to compare ctDNA yields obtained by four different manual extraction procedures, ii) to evaluate the addition of various carrier molecules...
into the plasma to improve ctDNA extraction recovery, and iii) to use next generation sequencing (NGS) technology to analyze KRAS, BRAF (B-raf proto-oncogene), and NRAS (neuroblastoma rat sarcoma viral oncogene homolog) somatic mutations in ctDNA from patients with metastatic colorectal cancer.

**Material and Methods**

**Subjects**

Venous blood with EDTA (9–10 mL) was obtained from thirty-two patients of the Department of Oncology, University Hospital in Hradec Králové who suffered from metastatic colorectal carcinoma. The experimental group consisted of 17 men and 15 women with a median age of 72 years (range 60–83 years). The diagnosis of metastatic disease was based on computer tomography examination. The standard clinical and histopathological classification of tumors was performed, including molecular analysis of KRAS, BRAF, and NRAS in formalin-fixed paraffin-embedded (FFPE) tumor tissue specimens. Histological verification of metastatic lesions was not required. The collections were performed with their informed consent.

**Extraction of ctDNA**

Within 1 h after collection, the blood specimens were centrifuged at 1300 g at 25 °C for 10 min; 2–3 mL supernatant (part I) was used for the preparation of pooled plasma. Consequently, 800 µL pooled plasma aliquots were spun at 12,000 g at 4 °C for 10 min. The 750 µL supernatant was transferred into a new plastic tube and stored at −20 °C. For ctDNA extraction, the following methods were used according to the manufacturer’s instructions adapted to a 750 µL plasma volume: QIAamp DNA Mini Kit (the spin protocol for DNA purification from blood or body fluids; Qiagen, Germany), QIAamp DSP Virus Spin Kit (Qiagen, Germany), NucleoSpin Plasma XS Kit (Macherey-Nagel, Germany), and Agencourt Genfind v2 Kit (Beckman Coulter, USA). The elution volume of TRIS-EDTA buffer was 35 µL. All extractions were performed in hexaplicates.

To evaluate the effectiveness of the used extraction procedures for short (<200 bp) and longer (200–500 bp) ctDNA fragments, 10 µL of GeneScan 500 LIZ Dye Size Standard (Applied Biosystems, UK) was added into another tube with the aliquot, and co-extracted along with ctDNA molecules. Then, fragmentation analysis in the ABI 3130 Genetic Analyzer (Life Technologies, USA) followed. The recovery of fragments was determined by their normalization to the longest (500 bp) fragment and expressed in percentages.

In second part of the study, the following amounts of carriers were added into the 750 µL pooled plasma aliquots: i) 17 µg carrier RNA (Qiagen, Germany), ii) 4 µg polyadenylic acid (poly(A); Roche Diagnostics, Germany), iii) 100 µg ultrapure glycogen (Invitrogen, USA), iv) 40 µg linear acrylamide (Invitrogen, USA), v) 40 ng ultrapure salmon sperm DNA (Invitrogen, USA), vi) 40 ng herring sperm DNA (Promega, USA). After that, the extraction process with the NucleoSpin Plasma XS kit was performed as described above. The extractions were carried out in hexaplicates, as well.

The third part of the study was focused on individual plasma specimens of the patients. From their residual 1–2 mL plasma (part II), a 750 µL supernatant was obtained as described above. After the addition of 4 µg polyadenylic acid, the extraction of ctDNA was carried out via the NucleoSpin Plasma XS kit. The extracts were stored at −80 °C until analysis.

**Analysis of ctDNA extracts**

Each extract was characterized by quantitative real-time PCR (Rotor-Gene 6000, Corbett Research, Australia) of the POLR2J housekeeping gene (gb Genetic Human DNA kit, Generi Biotech, Czech Republic). Using serial dilutions of Generi Biotech Standard Human Positive Control (20 ng/µL), a calibration curve ranging from 1 to 10,000 ng/mL was constructed, and ctDNA amounts in the extracts were determined.

For deep targeted NGS analysis, we used the Somatic 1 Master Kit (Multiplicom, Belgium) which enables molecular diagnostics based on multiplex BRAF, KRAS, and NRAS full exon amplifications (in total 30 amplicons with lengths of 168–255 bp) carried out in the MiSeq sequencing system (MiSeq Reagent Kit v2, 2x250 output; Illumina, USA). Analysis with amplicon-specific tagged primers was performed according to the manufacturer’s instructions with 7 µL of ctDNA extracts. As the wild-type control, 7 µL of ctDNA of two healthy subjects were included into each NGS run.

The presence of the mutations in ctDNA was validated by reverse hybridization strip assays: KRAS 12/13/61 StripAssay, BRAF StripAssay, and NRAS XL StripAssay (ViennaLab, Austria). The findings were finally compared with results of the FFPE tumor tissue analysis of the patients performed with the same technology in the frame of the routine diagnostic process. The established sensitivity of the strip technology for the variants was 1%.

**Bioinformatical analysis**

The secondary sequencing data analysis was initiated by generating raw binary base call files (BCL) from gray scale images of each cluster. For demultiplexing the samples, Illumina Miseq Reporter with a set up mismatch of 0 for each barcode was used. Paired FASTQ files were aligned to the reference Human genome HG19 by Burrows-Wheeler Algorithm (BWA) with the binary alignment map (BAM) output format. Variants were detected by Illumina Somatic Variant Caller Algorithm performed as a part of secondary analysis performed by MiSeq Reporter (MSR). The final variant calling format (VCF) files were annotated using an Illumina
Variant Studio online tool and visualized in the Integrative Genomics Viewer (IGV, Broad Institute of Massachusetts Institute of Technology and Harvard, USA). The detection threshold for mutations was set at 1%. The minimal read depth for detecting pathogenic variants was 100 bases at the given position.

### Statistical analysis

Concentrations of ctDNA were evaluated by using the Student t test. The normality of values was evaluated by the Shapiro-Wilk W test. Differences were considered to be statistically significant when \( P < 0.05 \).

### Results

The concentrations of ctDNA in pooled plasma extracts are demonstrated in Table 1. The highest levels of ctDNA were obtained by using the QIAamp DSP Virus Spin (the mean value 539 ng/mL) and NucleoSpin Plasma XS (448 ng/mL) kits. In these extracts, the ctDNA amounts were more than twofold higher than those from the QIAamp Mini Kit (163 ng/mL, \( P < 0.001 \)).

The procedures used differed in terms of the spectrum of DNA fragments extracted. The NucleoSpin method provided the highest extraction efficiency for fragments \(<200\) bp, in which most ctDNA molecules are contained. The Agencourt kit revealed a satisfactory recovery only for DNA fragments longer than 400 bp. Electroferograms for the NucleoSpin Plasma XS and QIAamp DSP Virus Spin fragments are demonstrated in Figure 1.

Figure 2 illustrates the influence of the carrier molecules added into plasma before the Nucleospin extraction procedure. Only the addition of 4 \( \mu \)g polyadenylic acid had a significant positive effect on the amount of ctDNA eluted (\( P < 0.05 \)). The other carriers had none (carrier RNA), or even a negative effect (glycogen, linear acrylamide, yeast tRNA, salmon or herring sperm DNA) on the ctDNA yield.

The Nucleospin technology, efficiency of which was enhanced by polyadenylic acid, was further used for individual ctDNA extractions from the thirty-two plasma specimens of the patients. In the extracts, the levels of ctDNA ranged from 50 to 580 ng/mL with a median value of 260 ng/mL.

---

**Tab. 1:** Concentrations of ctDNA and recoveries of fragments differing in size.

<table>
<thead>
<tr>
<th>Extraction</th>
<th>DNA concentration Mean (SD) ng/mL</th>
<th>Recovery of fragments*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>75/500 bp</td>
</tr>
<tr>
<td>QIAamp Mini</td>
<td>163 (24)</td>
<td>32</td>
</tr>
<tr>
<td>Nucleospin</td>
<td>448 (48)**</td>
<td>77</td>
</tr>
<tr>
<td>DSP Virus</td>
<td>539 (154)**</td>
<td>45</td>
</tr>
<tr>
<td>Agencourt</td>
<td>223 (69)</td>
<td>0</td>
</tr>
</tbody>
</table>

* 500 bp fragments were used as referent; SD standard deviation; ** \( P < 0.001 \)

---

![Fig. 1: Extraction recoveries of DNA fragments of GeneScan 500 LIZ Dye Size Standard when added into human plasma and extracted with the QIAamp DSP Virus Spin Kit (upper part) and the NucleoSpin Plasma XS Kit (lower part). Lengths of fragments: 75, 100, 139, 150, 160, 200, 250, 300, 340, 350, 400, 450, 490 and 500 bp. F is fluorescence. The latter extraction procedure showed lower losses of the fragments in the range of 75–200 bp.](image1)

![Fig. 2: Concentrations of ctDNA in the NucleoSpin extracts in relation to the used carrier molecule. DNA concentrations are expressed as means and standard deviations; * \( P < 0.05 \), ** \( P < 0.01 \), *** \( P < 0.002 \), **** \( P < 0.001 \).](image2)
The following NGS analysis was successful in all cases; 100% bases in the exons were sequenced bi-directionally. We achieved NGS run metrics as follows: cluster density 678 K/mm²; total number of reads 14.1 millions; pass filter reads 13.2 millions; mean number of reads 0.5 million per sample; average amplicon coverage 15,000-fold; and the total amount of basecalls >Q30 was 86.0%.

The sequencing data revealed five cases (16%) of ctDNA mutated in KRAS. Three patients were mutated in KRAS codon 13 (the mutation c.38G>T, p.Gly13Asp, rs112445441 in all three specimens), one in KRAS codon 12 (c.35G>T, p.Gly12Val, rs121913529), and another one in KRAS codon 61 (c.182A>G, p.Gln61Arg, rs121913240). In BRAF, one subject (3%) was found to have p.Val600Glu activating mutation (c.1799T>A, rs113488022). No mutations in the NRAS gene were obvious. The results of the ctDNA NGS analysis agreed with those obtained by the reverse hybridization technique, and with the data on FFPE, except one subject with tumor tissue mutated in KRAS codon 12 (c.34G>A, p.Gly12Val, rs121913530) and negative in the corresponding ctDNA specimen. Thus, an agreement of 86% was found between tumor tissues and ctDNA.

Discussion

Knowledge of genetic background helps in selecting an individual approach to metastatic colorectal cancer treatment, including targeted biological therapy. However, primary tumor tissue is not always available, and could be of insufficient quality, or could have been obtained a long time before the metastases were diagnosed. Moreover, several reports have indicated the status of somatic mutations in metastases changes in the course of therapy as a result of tumor heterogeneity, clonal expansion, and selection (16, 17). These changes are responsible for acquired resistance developing within a few months (18). Since invasive and painful biopsies of metastatic tissue are often difficult to obtain, ctDNA testing, available at any disease stage, seems to be a good alternative for analyzing mutations during the follow-up period.

One aim of this study was to increase the efficiency of the ctDNA extraction process, when using 2–3 mL of blood or 750 µL of plasma, respectively, is used. Larger blood collections during the follow-up period of metastatic patients are sometimes difficult to obtain. There are a lot of manufacturers providing commercial products for cfDNA extraction and purification. We examined two of them based on the manual spin technology with QIAamp DSP Virus Spin Kit or without the addition of carrier molecules (NucleoSpin Plasma XS Kit), and with a protocol that uses paramagnetic separation beads (Agencourt Genfind v2 Kit). The results were compared to the QIAamp DNA Mini Kit, a universal and robust extraction product used in clinical labs for over fifteen years. For the evaluation, quantitative real-time PCR analysis was preferred to the spectrophotometric and fluorometric measurements that often result in interference of carrier polynucleotide chains and overestimation of low-copy DNA molecules.

Mouliere et al. reported that more than 80% of ctDNA fragments in the plasma of metastatic patients were shorter than 145 bp, with a large proportion of the ctDNA fragments <100 bp (19, 20). Our results revealed that the NucleoSpin Plasma XS process is highly effective for ctDNA fragments in the size range of 75–200 bp, despite the fact that no carrier molecules are used in it. The mean concentration of extracted ctDNA was 448 ng/mL. It corresponded to almost 68,000 GE/mL and agreed with previous studies (7, 19–21).

Next, we studied how the addition of various carrier molecules effects the extraction efficiency. The QIAamp DNA Mini Kit instructions recommend using carrier DNA for extractions of low copy number DNA (<10,000 GE/mL), although the carrier DNA is not included in the kit. In other Qiagen products, carrier RNA molecules are included and used regardless of the ctDNA amounts expected in the specimen. Shaw et al. previously showed that for maximum improvement of the DNA yield, the ideal ratio of carrier RNA to DNA was between 10:1 and 50:1; ratios outside this range do not enhance DNA recovery as successfully (22). We added 4 µg of polyadenylic acid into 750 µL of plasma (ratio 12:1) before the NucleoSpin extraction process, thus increasing the ctDNA yield in eluates.

The addition of poly(A) carrier RNA but not glycogen previously increased the recovery of automated silica-based extractions (BioRobots EZ1 and BioRobots M48, Qiagen) by an average of 24% (23). In another paper, a five-fold increased recovery was obtained in DNA extractions carried out on silica-based monoliths within a microfluidic device when poly(A) carrier RNA was added to the chaotropic salt solution (22). However, the Qiagen carrier RNA added into the plasma specimens of our patients in a ratio of 50:1 (RNA:ctDNA) had no effect on the NucleoSpin ctDNA recovery. Not only the proper carrier RNA:ctDNA ratio but also the length of poly(A) chains and their folding in space probably play an important role in the extraction process.

The remaining carriers reduced the final ctDNA amounts. Cheung et al. reported that the addition of yeast tRNA or salmon sperm DNA prior to purification by silica particles resulted in significantly decreased recovery of HCV RNA from sera (24). Thus, glycogen, linear acrylamide, yeast tRNA, salmon sperm DNA, or herring sperm DNA, are not suitable substances to improve yields of commercial silica-based extraction procedures. On the other hand, when used as co-precipitants, they can facilitate recovery of the target DNA molecules in the phenol/chloroform extraction from eukaryotic or prokaryotic cells (25).

For NGS, we used a Multiplicom Somatic 1 Master Kit manufactured for FFPE tumor tissue DNA analysis. To our best knowledge, this is the first study that uses the kit for ctDNA testing. We took into account similar properties of FFPE DNA and ctDNA. Clear and sensitive results were obtained from all the tested ctDNA specimens. The total number of plasma ctDNA mutated in KRAS or BRAF
reached 19%, the concordance of the ctDNA with FFPE results was 86%. Similar discrepancies have been previously reported (14, 26, 27). A lower than expected frequency of the mutated tissue specimens (28) could be explained by the small number of specimens in the study, by the elevated mortality rate of the subjects with more aggressive types of mutations, or by other reasons. No mutations were found in the NRAS gene. This finding corresponds to the generally low prevalence of NRAS mutations in colon tumors (29).

A lot of somatic mutations exist in other genes (PTEN, EGFR, PIK3CA, ERBB2, PIK3R1, etc.) in colorectal cancer. Although commercial kits for NGS analysis are currently available, their clinical use for predictive testing is not yet obligatory. Similarly, the clinical importance of determining these mutations in ctDNA has not yet been demonstrated.

Conclusions

In conclusion, testing somatic mutations in ctDNA seems to be a promising tool to monitor dynamically changing genotypes of tumor cells circulating in the body, and causing disease relapse. The optimized process of ctDNA extraction should help to obtain more reliable data on KRAS, BRAF, NRAS, and several other genes when using NGS and/or other molecular techniques in patients with metastatic colorectal cancer. We believe that our work would contribute to better standardization of the pre-analytical phase of ctDNA analysis, and to a broader use of ctDNA in clinical practice.

Acknowledgements

The study was supported by the project MZ CR – RVO (FNHK, 00179906) of the Ministry of Health, Czech Republic, and by the projects PRVOUK P37/11 and P37/09 of the Charles University in Prague, Faculty of Medicine in Hradec Králové, Czech Republic.

References


Received: 10/02/2016
Accepted: 18/04/2016