

Original Manuscript

Mutational landscape of plasma cell-free DNA identifies molecular features associated with therapeutic response in patients with colon cancer. A pilot study

Klara Cervena^{1,2,✉}, Barbara Pardini^{3,4}, Marketa Urbanova^{1,2},
Sona Vodenkova^{1,5}, Pazourkova Eva^{2,6}, Veronika Veskrnova⁷,
Miroslav Levy⁸, Tomas Buchler⁷, Martin Mokrejs^{1,5}, Alessio Naccarati^{3,4,✉},
Pavel Vodicka^{1,2,5} and Veronika Vymetalkova^{*,1,2,5,✉}

¹Department of Molecular Biology of Cancer, Institute of Experimental Medicine of the Czech Academy of Sciences, Videnska 1083, 142 20 Prague, Czech Republic, ²Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Albertov 4, 128 00 Prague, Czech Republic, ³IIGM Italian Institute for Genomic Medicine, Sp142 Km3.95, 100 60 Candiolo, Turin, Italy, ⁴Candiolo Cancer Institute, FPO-IRCCS, Sp142 Km3.95, 100 60 Candiolo, Turin, Italy, ⁵Biomedical Centre, Faculty of Medicine in Pilsen, Charles University, Alej Svobody 76, 323 00 Pilsen, Czech Republic, ⁶Department of Nephrology, First Faculty of Medicine, Charles University and General Faculty Hospital in Prague, U nemocnice 2, 121 08 Prague, Czech Republic, ⁷Department of Oncology, First Faculty of Medicine, Charles University and Thomayer Hospital, Videnska 800, 140 59 Prague, Czech Republic and ⁸Department of Surgery, Thomayer University Hospital, Videnska 800, 140 59 Prague, Czech Republic

*To whom correspondence should be addressed. Tel: +420 2 296 4 2699; Fax: +420 2 410 6 2782; Email: veronika.vymetalkova@iem.cas.cz

Received 10 March 2021; Editorial decision 29 June 2021; Accepted 30 June 2021.

Abstract

Cell-free DNA (cfDNA) has recently been used as a non-invasive diagnostic tool for detecting tumour-specific mutations. cfDNA may also be used for monitoring disease progression and treatment response, but so far researchers focused on one or few genes only. A genomic profile may provide better information on patient prognosis compared to single specific mutations.

In this hypothesis-generating study, we profiled by whole exome sequencing serial plasma samples from 10 colon cancer (CC) patients collected before and after 5-fluorouracil-based therapy, and one year after diagnosis to determine alterations associated with treatment response. In parallel, genome profiling was also performed in patients' corresponding tumour tissue to ascertain the molecular landscape of resistant tumours.

The mutation concordance between cfDNA and tumour tissue DNA was higher in more advanced tumour stages than in the early stages of the disease. In non-responders, a specific mutation profile was observed in tumour tissues (*TPSD1* p.Ala92Thr, *CPAMD8* p.Arg341Gln, *OBP2A* p.ArgTyr123CysHis). A pathogenic *APC* mutation (p.Ser1315Ter) was detected only in cfDNA of one poor responder one year after the diagnosis and after therapy termination. Another poor responder presented a likely pathogenic *TP53* mutation (p.Arg110Pro) in cfDNA of all plasma samplings and in tumour tissue.

In conclusion, cfDNA could be used for genetic characterisation of CC patients and might be clinically useful for non-invasive therapy response monitoring.

Introduction

Colorectal cancer is the third most commonly diagnosed malignancy and the second leading cause of cancer-related deaths worldwide (1, 2), with colon cancer (CC) representing 60–70% of all cases (3). The five-year survival rate for colorectal cancer patients in stage I or II reaches up to 90%, whereas it is less than 10% for patients in stage IV. For decades, fluoropyrimidine-based chemotherapy has been the first choice of CC treatment (4). Due to the obvious heterogeneity in disease course, significant efforts are currently dedicated to discriminate patients who will benefit from chemotherapy. Different responses to therapies, particularly chemoresistance, can be conferred by genetic or epigenetic modifications with consequent alteration of drug sensitivity. Despite the encouraging progress in CC therapy, to date patient response rates to treatment remain low and the benefit from fluoropyrimidine-based therapies is often limited by the development of chemoresistance (5).

There is currently no gold standard method to differentiate therapy responders from non-responders, and no satisfactory molecular markers have been identified yet (6). In most of the cases, tumour tissue is the preferred specimen source for accurate and comprehensive genomic profiling of cancer patients in relation to therapy response and resistance. However, occasionally, tumour tissues may be unavailable or limited, or the patient could not be able to undergo an invasive biopsy due to the high risk of the procedure or comorbidities. Moreover, a tissue sample represents a single snapshot in time and is subjected to selection bias due to tumour heterogeneity. For these reasons, non-invasive liquid biopsies based on the detection of cell-free DNA (cfDNA) in the blood offer a useful alternative for such analyses (7). Recently, Zhou et al. (8) proved that cfDNA from metastatic colorectal cancer patients undergoing treatment with bevacizumab helped to identify recurrent somatic copy number alterations of clinical relevance in the *POLR1D* gene. It is essential to have reliable data to demonstrate how precisely a liquid biopsy may substitute the genomic profiling of tumours (9), as well as to understand its appropriate clinical settings and limitations.

Tumour-specific DNA mutations can be detected in circulating tumour DNA, which represents a component of the cfDNA. Extensive studies have demonstrated that it could be used for the early diagnosis of a variety of cancers (as reviewed in (10–14)), including breast (15, 16), pancreatic (17, 18), lung (19), and colorectal cancer (20, 21). The analysis of circulating tumour DNA has started to spread in clinics since it features higher sensitivity and specificity than other non-invasive traditional biomarkers such as faecal occult blood test for colorectal cancer, or circulating carcinoembryonic antigen (CEA) for lung, breast, liver, or pancreatic cancer (14, 22, 23). In this sense, liquid biopsy is also becoming more widely used as a tool for identifying genomic alterations to assess prognosis, guide therapy and monitor residual disease (24). Together with the recent advances in DNA sequencing technology and ‘omics’-based approaches, liquid biopsy promises a new strategy for the development of novel molecular biomarkers in CC disease.

Considering that the surgical tumour resection with chemotherapy is the only choice for CC treatment, efforts are needed to elucidate the clinical value of cfDNA in CC regarding optimal treatment strategies and benefits.

In the current hypothesis-generating study, for the first time, we prospectively analysed plasma cfDNA from repeated samples (peri- and post-operative) and its capability in CC patients in predicting the response to therapy among good responders and non-responders. In addition, we evaluated the tumour genome of the same patients to

identify mutations associated with therapy response and compared the results with cfDNA.

Materials and Methods

Sample collection

In this prospective and single-centre study, newly diagnosed histologically confirmed CC patients were recruited by the Department of Surgery at the Thomayer Hospital (Prague, Czech Republic). In order to minimise the heterogeneity of the cohort, the following inclusion criteria were applied: (i) newly diagnosed sporadic form of CC without neoadjuvant (i.e. pre-operative) chemoradiotherapy and (ii) adjuvant treatment based on 5-fluoruracil (5-FU: consisting of either a bolus 5-FU regimen, an infusional 5-FU regimen, capecitabine, or a fluoropyrimidine in combination with oxaliplatin). Patients with any personal history of previous malignancy, or with CC-associated well-defined inherited syndromes (including Lynch syndrome, familial adenomatous and MUTYH-associated polyposis) were excluded from the study. In total, 10 CC patients were included in the study. For each patient, tumour tissue pairs were collected together with blood plasma at the time of surgery (blood sampling always preceded the tissue collection with a time lapse ranging from 1 to 30 days between blood and tissue sampling according to the pre-hospitalisation appointment). Subsequently, plasma samples were collected from the same patients every 6 months for a total of three samplings; this was to reflect the condition at the diagnosis (pre-operative sampling), at the termination of the planned therapy and after additional 6 months (post-operative samplings). For one CC patient, monitoring was exceptionally carried out up to 18 months after diagnosis.

All tumour tissues were snap-frozen immediately after surgical resection, stored at -80°C , and subsequently used for experimental analyses.

Disease characteristics and treatment data, such as tumour-node-metastasis (TNM) stage, histopathological grade, and the type of chemotherapy regimen received were collected after surgical resection. All the information regarding response to therapy and survival were also collected, including disease progression, relapse, distant metastasis, date of the last examination/death and details about the adjuvant chemotherapy adverse effects. Good therapy responders were defined as patients who did not require a change of chemotherapy regimen, were in remission after the prescribed number of cycles, and did not suffer from any adverse effects caused by chemotherapy. The last update of patients’ follow-up for this study was done in May 2020.

All participants were informed about all aspects of the study, agreed with the study purpose and procedures to be undertaken and provided informed consent. The Ethics Committee of the participating hospital (Ethical Committee for Institute for clinical and experimental medicine and Thomayer hospital, ID of approval G-10-06-38) approved the protocol of the present study which adhered to the ethical guidelines as set out in the Helsinki Declaration.

Extraction of plasma cfDNA

Three serial blood samples were collected at the following defined time points: initial pre- (within 1–30 days before surgery) and post-surgery (6- and 12-month intervals after surgery). For genomic profiling of cfDNA, two EDTA tubes (9 ml each) of whole peripheral blood were collected. Blood was spun within 1 h after its collection at 1500g for 20 min, and the plasma fraction (up to 2 ml) was frozen at -80°C . Thawed plasma samples were centrifuged once

more at 20 000g for 10 min to separate any further cellular portions. cfDNA was extracted from aliquots of plasma using the QIAamp circulating nucleic acid kit (Qiagen) according to the manufacturer's protocol. cfDNA was eluted twice through each column to maximise yield, and finally stored at -20°C in 20 μl aliquots in DNA LoBind tubes (Eppendorf). The quantification of cfDNA for each plasma sample was determined by Bioanalyzer 2100 instrument with a High Sensitivity DNA kit (Agilent Technologies) and by Qubit fluorimeter (High Sensitivity DNA assay, Qubit 3.0, Thermo Fisher).

Extraction of DNA from tissues

Tumour samples, stored at -80°C , were homogenised by MagNAlyser Instrument (Hoffmann-La Roche). DNA from tumour sections was extracted using DNA Allprep kit (Qiagen) according to the manufacturer's instructions and stored at -20°C . DNA concentration was measured with Qubit fluorimeter (DNA Broad Range assay, Qubit 3.0, Thermo Fisher).

Enrichment of DNA spanning exonic sequences from plasma cfDNA

We enriched protein-coding DNA using Agilent Human Whole Exome kit v6 and v7 according to manufacturer's instructions.

Library preparation and sequencing of plasma cfDNA

Fragments of 200–400 base pairs (bp) were selected using Agencourt AMPureXP beads (Beckman Coulter, Brea, CA, USA). The different libraries were prepared using the Rubicon ThruPLEX or Agilent SureSelect XT HS Library Preparation protocols following the manufacturer's instructions and starting with 10–60 ng of cfDNA as input material. Library quantification and quality control were performed using the KAPA Library Quantification Kit (Illumina) and Bioanalyzer 2100 (Agilent) with a High Sensitivity DNA Kit (Agilent). In total, 19 cfDNA samples originating from 10 patients were sequenced on three Illumina HiSeqflow cells. Seven samples were pooled together for sequencing in one flowcell lane on Illumina HiSeq2500 using 2×125 nt long reads. Twelve samples were sequenced on HiSeq4000 (Illumina) in paired-end mode yielding 2×100 nt reads. The median coverage was 32 \times (per-sample median value was calculated as a median of medians calculated for every region covered by the exome enrichment probes).

Sequencing of tumour DNA

Ten DNA samples from tumour tissues were sheared using the CovarisM220 ultrasonicator. Exome enrichment libraries were prepared using the SureSelect XT HS Library Preparation and SSELXT2 Human All Exon V7 Kits (both Agilent) according to the manufacturer's instructions in a similar way as for cfDNA. All samples were pooled together and DNA fragments of 400 nt were sequenced using paired-end mode yielding 2×100 nt long reads on a HiSeq4000 (Illumina). The median coverage was 36 \times (individual, per-sample median value was calculated as a median of medians calculated for every region covered by the exome enrichment probes).

Analysis of sequencing data

The quality of raw reads was evaluated using FastQC (v0.11.5) and MultiQC (1.0.dev0) and reads that did not meet the defined standards (Quality Phred score cutoff: 25, Minimum required sequence length >34 nt, maximum trimming error rate 0.1, clipping Illumina

adaptor sequences, trimming first 12–22nt) were removed, trimmed or edited using FASTX Toolkit (v0.0.14) and Cutadapt (v1.9.1).

Alignment to the human genome reference sequence (hg19 without decoy sequences) was performed using Burrows-Wheeler Aligner (BWA, v0.7.12-r1039) in default settings for all whole exome sequencing (WES) data, including plasma samples and tumour tissues. PCR duplicates were removed using PicardTools (v2.4.1) with optical distance parameter set to 100 (the default).

Variants were discovered based on Bayesian methods for inferring haplotypes using FreeBayes (v1.0.2-29-g41c1313). Variants were filtered out according to the following criteria: Phred-scaled quality score (less than 20), Phred-scaled quality score/alternate allele observations, with partial observations recorded fractionally (less than 10), alternate observations on the forward strand (>0), alternate observations on the reverse strand (>0), reads supporting the alternate balanced to the left ($5'$) of the alternate allele (less than 1 read), reads supporting the alternate balanced to the right ($3'$) of the alternate allele (less than 1), and total read depth at the locus (less than 109 and 219). Variants were annotated, and their effect was predicted using SnpEff (4.3t) based on Human genome database UCSC hg19, (http://downloads.sourceforge.net/project/snpeff/databases/v4_3/snpeff_v4_3_hg19.zip) and ACMG (American College of Medical Genetics) classification, all characterised by Varsome (<https://varsome.com/>, (25)). Further, Gnomad 2.1.1 and Genomics Data Commons available on VarSome. Each identified variant that was classified in Varsome as pathogenic, likely pathogenic, or variants with uncertain significance (VUS) was further studied in Clinvar (<https://www.ncbi.nlm.nih.gov/clinvar/>), a tool that aggregates information about genomic variation and its relationship to human health.

Comparison of mutations between plasma and tumour samples

For tumour/plasma comparison, we identified all variants called in data from concomitant plasma and tumour samples, as described above. We retained all variants adequately covered in both samples (minimum 30 reads in the tumour, minimum 10 reads in synchronous plasma data).

Results

Patients' characteristics

The clinical characteristics of patients included in the study are described in Table 1. Patients were mostly males (70%), and the mean age was 66.4 ± 5.9 years. In terms of therapy, no subjects received chemotherapy or radiation prior to surgical resection or blood sampling. The majority of patients enrolled in the study at the time of surgery showed an advanced CC stage (80% were stages III–IV). All patients received chemotherapy through different fluoropyrimidine-based regimes: a fluoropyrimidine either in monotherapy or a combination with leucovorin (seven patients) and three patients received a combination of fluoropyrimidine with oxaliplatin—FOLFOX or CapOX. Four patients died during follow up, from which two had a tumour recurrence.

Analysis of sequencing data

After alignment, a range from 4.6 to 76.7 million of properly aligned read pairs per sample were obtained. In summary, the percentage of properly mapped read pairs reached 99% for all samples. The average read depth of analysed genes within each and total read

Table 1. Patients' clinical characteristics

		N = 10
Age	Years ± SD	66.4 ± 5.9
Gender	Men	7
	Women	3
CC tumour stage ^a	II	2
	III	4
	IV	4
Tumour size (T) ^a	1	0
	2	0
	3	7
	4	3
Lymph node status (N) ^a	0	6
	>1	4
Metastatic status (M) ^a	0	6
	>1	4
Grade ^a	2	6
	3	4
MSI	MSS	8
	MSI	2
Presence of a local recidive ^a	No	8
	Yes	2
5-FU based therapy ^a	No	0
	Yes	10
Response to therapy	Good	3
	Poor	7
Living status at follow up	Dead	4
	Alive	6

^aAccording to the International Union Against Cancer (UICC) tumour-node-metastasis (TNM) stage system.

Abbreviations: CC, colon cancer; MSI, microsatellite instability; MSS, microsatellite stable; MSI, microsatellite instable; SD, standard deviation; 5-FU, 5-fluorouracil.

depth for each patient can be found in [Supplementary Table S1](#) and [Supplementary Table S2](#), respectively.

Mutational burden concordance between cfDNA and tumour DNA

The mutational concordance between cfDNA fragments and genomic DNA from colon tumour was assessed. We initially focused on those genes commonly mutated in colorectal cancer, and their specific mutations (as listed in [Tables 2–4](#) and [Supplementary Tables S3](#)) whose altered products are known to correlate with chemotherapy efficacy (by OMIM, www.omim.org; <http://www.mygenomics.com/cancer-panels-gene-list/>). WES in tumour DNA was performed in order to identify somatic tumour-specific single nucleotide variants (SNV) and short insertion/deletion (Indel) mutations characterising colon tumours. At least one SNV (either pathogenic, likely pathogenic, or a variant of uncertain significance (VUS)) was identified in tumour DNA of all patients. The median number of mutations for the selected genes was 7.5 (range 3–16). The most frequent somatic mutations at the time of enrolment were in *TP53* (70%) and *APC* (60%) genes; however, in two patients no mutations in these two genes were detected. The number of somatic mutations tended to be higher in stage IV patients ([Tables 2–4](#) and [Supplementary Table S3](#)).

In all patients, at least one mutation in the enriched genes was detected in preoperative plasma, which was consistent with the paired tumour tissue. However, several discrepancies between the genomic landscape of cfDNA and tumour DNA were observed. These inconsistencies were predominantly noticed for CC stages II and III.

Among these patients, in fact, several mutations observed in tumour DNA were not detected in cfDNA, in particular those in *APC* (4 patients), *TP53* (4 patients), *SMAD4* (3 patients), and *KRAS* (2 patients) genes. As expected, with the increase of TNM stage, mutation concordance between cfDNA and tumour DNA increased. On the other hand, one stage IV patient carried a mutation in the *DCC* gene in cfDNA which was not present in the tumour DNA.

Genomic landscape of tumour DNA in relation to therapy response

In the clinical follow up, we identified three good responders (those who benefit from the chemotherapy, with no relapses and having a complete response with no residual cancer) and seven non-responders (patients who lack any therapy response, die very early after diagnosis) ([Table 1](#)). The WES data from tumour DNA derived from good and non-responders were compared. In total, 71 705 SNVs and Indels have been identified as different between these two groups: 1555 with high, 23 385 with moderate and 23 471 with putative modifier effect. No new variants were detected either among good responders or non-responders.

The identified variants were further filtered according to the presence of the genetic variation in at least four non-responders and in none of the responders. In total, 623 SNVs or Indels were detected; after stratification for putative effect, 404 variants with high or moderate effect were collected. The most pronounced difference between the two groups of patients was observed for the *GPR50* gene (c.1594A>G and c.1505_1516delCCACTGGCCACA). Except for one non-responder, all the others carried a homozygous variant genotype for both variants of this gene when compared to good responders. After classification for the putative effect, both variants were classified as moderate and benign. Additionally, all non-responders carried either a variant (c.1509T>A) in homozygosis or heterozygosis of the *THSD7B* gene that was absent in good responders. After classification for the putative effect, this variant was classified as moderate and benign similarly to those in *GPR50*.

According to the presence of the genetic variation in at least five non-responders and in none of the responders, differences were noticed for *TPSD1* c.274G>A (p.Ala92Thr), *CPAMD8* c.1022G>A (p.Arg341Gln), and *OBP2A* c.367_370delCGCTinsTGCC (p.ArgTyr123CysHis). After a classification by ACMG by Varsome, these three variants were characterised as of uncertain significance (VUS).

Genomic landscape of cfDNA between good and non-responders to therapy

Due to the limited insights available based on tumour-tissue datasets, we focused on the dynamic changes in cfDNA samples to identify mutations attributable to 5-FU resistance. In total, 106 132 SNVs and Indels differed between good responders and non-responders in cfDNA profiling: 2254 with high, 31 300 with moderate and 32 959 with putative modifier effect. No new variants were found in good responders compared to non-responders and vice versa, as already observed for tumour DNA.

Called variants were filtered according to the presence of the variation in at least four non-responders and in none of the responders. Following these criteria, 3925 SNVs or Indels were selected and after stratification for putative effect, a total of 2699 variants resulted with high, moderate or putative modifier effects. The highest difference between these two groups of patients was observed for the *ABCD1* gene (c.*8G>C). Except for one non-responder who

Table 2. Summary of pathogenic, likely pathogenic mutations and VUS in tumour and ctDNA samples of stage II

Stage	II						
	T	cfDNA/I	III	IV	T	cfDNA/I	II
Sample ID	220				392		
Sex	Female				Female		
Age	62				69		
MSI	MSS				MSS		
Therapy response status	Poor				Good		
<i>ACVR1B</i>							
<i>APC</i>	c.3340C>T c.4033G>T						
<i>CDC27</i>							
<i>DCC</i>							
<i>EGFR</i>	c.2088A>G	c.2088A>G	c.2088A>G	c.2088A>G			
<i>EP300</i>							
<i>ERBB2</i>	c.528C>T	c.528C>T	c.528C>T	c.528C>T			
<i>FBXW7</i>							
<i>KRAS</i>							
<i>MAP7</i>							
<i>MLH1</i>							
<i>MSH3</i>							
<i>MUTYH</i>							
<i>MYC</i>							
<i>MYO1B</i>							
<i>PIK3CA</i>							
<i>PIK3R1</i>							
<i>POLE</i>	c.4523G>A	c.4523G>A	c.4523G>A	c.4523G>A			
<i>SMAD2</i>							
<i>SMAD4</i>							
<i>SMAD7</i>	c.115G>A	c.115G>A	c.115G>A	c.115G>A	c.115G>A	c.115G>A	c.115G>A
<i>TCERG1</i>	c.409G>A	c.409G>A	c.409G>A	c.409G>A			
<i>TCF7L2</i>							
<i>TGFBR2</i>							
<i>TP53</i>					c.586C>T		

MSS, microsatellite stable; MSI, microsatellite instable; T, tumour tissue; cfDNA, cell-free DNA.

ID/I represents the sampling at the time of the diagnosis.

ID/II represents the sampling 6 months after diagnosis.

ID/III represents the sampling one year from the diagnosis.

ID/IV represents the sampling 18 months since the diagnosis.

Pathogenic mutations are in red.

Likely pathogenic mutations are in green.

VUS are in blue.

* Due to long notation, detailed information in [Supplementary Table S3](#)

was heterozygous for this variation, all non-responders carried the homozygous variant genotype when compared to good responders. After classification for the putative effect, this variation was classified as benign.

According to the presence of the variation in at least five non-responders, the absence of this variation in good responders, and after classification of variants by ACMG by Varsome, only two variants of uncertain significance were noticed: *CPAMD8* c.1022G>A (p.Arg341Gln) and *OBP2A* c.367_370delCGCTinsTGCC. With less stringent criteria of difference between the group of responders (4 heterozygous genotypes and more), additional three VUS were identified: *SPANXD* c.175_177delTTTinsGTG (p.ArgTyr123CysHis), *CHMP4A* c.586G>A (p.Gly196Arg), *OR4N2* c.171_172delCCinsAG (p.Pro58Ala).

On the other hand, all the good responders presented an insertion classified as VUS in the *ZXDB* gene (c.364_368delGAGG AinsAAGGC; p.GluGlu122LysAla), which was not carried from non-responders.

Monitoring cfDNA in serial plasma samplings before and after 5-FU-based therapy

For seven patients, cfDNA was isolated from plasma samples collected before and after chemotherapy to monitor the changes in the mutation patterns ([Supplementary Table S3](#)).

Interestingly, patient ID221, a non-responder, showed a pathogenic *APC* mutation (c.3944C>A; p.Ser1315Ter) in cfDNA in its third sampling (i.e. 1 year after surgery). This particular mutation was not found in the same patient either in the tumour DNA or in the cfDNA collected at the time of diagnosis. This mutation could be a signal of change that could help in clarifying the poor response to therapy. *APC* is a known tumour suppressor gene that acts as an antagonist of the Wnt signalling pathway and is involved in cell migration and adhesion, transcriptional activation, and apoptosis.

Another non-responder, ID327, presented a likely pathogenic *TP53* mutation (c.329G>C; p.Arg110Pro; rs11540654) in all the repeated plasma samples that was also present in tumour DNA. This

Table 4. Summary of pathogenic, likely pathogenic mutations and VUS in tumour and ctDNA samples of stage IV

Stage	IV			
Sample ID	221	402	464	481
Sex	Male	Male	Male	Male
Age	60	66	67	58
MSI	MSI-L	MSS	MSS	MSS
Therapy response status	Poor	Poor	Poor	Poor
	T	T	T	T
	cfDNA/I	cfDNA/I	cfDNA/I	cfDNA/I
ACVR1B				
APC		c.3944C>A	c.4393_4394del*	c.712C>T
CDC27			c.714delT c.701_704del*	c.714delT c.701_704del*
DCC	c.1256A>G	c.1256A>G		c.714delT c.701_704del*
EGFR				
EP300				
ERBB2				
FBXW7				
KRAS		c.183A>T	c.183A>T	
MAP7				
MLH1				
MSH3				
MUTYH		c.157+75G>A	c.157+75G>A	c.195_203dup*
MYC				
MYO1B	c.1883G>A	c.1883G>A		
PIK3CA		c.1624G>A	c.1624G>A	
PIK3R1				
POLE				
SMAD2				
SMAD4				
SMAD7				
TCERG1				c.115G>A
TCF7L2		c.1484C>G	c.1484C>G	
TGFBR2		c.742C>T	c.742C>T	c.916C>T
TP53			c.817C>T	c.817C>T

MSS, microsatellite stable; MSI, microsatellite unstable; T, tumour tissue; cfDNA, cell-free DNA.

ID/I represents the sampling at the time of the diagnosis.

ID/II represents the sampling 6 months after diagnosis.

ID/III represents the sampling one year from the diagnosis.

ID/IV represents the sampling 18 months since the diagnosis.

Pathogenic mutations are in red.

Likely pathogenic mutations are in green.

VUS are in blue.

*Due to long notation, detailed information in [Supplementary Table S3](#).

could be another example of poor response to therapy due to a persistence of a dysfunctional tumour-derived mutated *TP53* protein.

Discussion

In the present hypothesis-generating study, we sequenced plasma cfDNA from CC patients to investigate the tumour genome evolution during the chemotherapy. To the best of our knowledge, this is the first study based on WES data that compared the mutational status in colon tumour tissues with those of serial plasma cfDNA samplings in association with the response to therapy.

Overall, we observed a higher mutation detection rate in plasma cfDNA and tumour DNA of patients with metastatic CC in comparison to stage II and III patients. This is in agreement with previous studies (21, 26) and confirms that WES is a feasible approach for liquid biopsy in patients with advanced stages of CC. On the other hand, we have also observed several differences between the genomic landscape of cfDNA at the time of diagnosis and the corresponding DNA profile from tumour tissues. These inconsistencies were predominantly noticed for stages II and III. Among these patients, several mutations found in tumour DNA were not detected in cfDNA, in particular in *APC* (four patients), *TP53* (four patients), *SMAD4* (three patients) and *KRAS* (two patients) genes. As expected, with the increase of TNM stage, the mutation concordance between cfDNA and tumour DNA increased. This is in agreement with other studies where, based on mutation spectrum in primary tumour, the detection of mutations in plasma cfDNA was significantly associated with disease stage: 24.0% in stage I, 45.0% stage II, 27.3% stage III and 87.5% stage IV (27). Conversely, such as the case of a patient with CC IV stage in our study, a mutation of the *DCC* gene in cfDNA was not found in the corresponding tumour DNA. This specific discrepancy between specimen in CC IV stage patient could be explained by the fact that mutant DNA from not only primary tumour but also from distant metastases within a patient can be sampled by the cfDNA approach. The liquid biopsy approach may thus reduce the risk of missing a mutation due to tumour heterogeneity or sampling of primary tissue only.

Tumour heterogeneity is one of the major causes of variable response to treatment in CC patients. More specifically, the presence of multiple subclones bearing different molecular and phenotypic traits within the same tumour, has been reported in a number of cancers, including CC (28). Researchers have shown the effect of intratumour heterogeneity on the accuracy of diagnosis, choice of therapy and mechanism of resistance. Previous studies have reported the concordance of genetic profiles between cfDNA and DNA from tumour or metastatic tissues (29). Sottoriva et al. (30) suggested a 'big bang model' hypothesis, in which mutations responsible for tumour development and progression occur early in CC and thus biological tumour performance is determined early ('born to be bad'). However, CC harbours on average about 80 mutations and yet fewer than 15 of them seems to be the driver force for tumorigenesis and progression (31). In agreement with previous statements, Wood et al. (32) hypothesised that patient's prognosis is most probably determined by the overall mutation landscape and aberrant pathways more than any single specific mutation.

To determine alterations associated with treatment response, we have focused on finding differences between good responders and non-responders with a specific workflow (Figure 1). First, we compared the tumour DNA between the two groups, then we analysed the plasma cfDNA at the time of surgery and finally the serial cfDNA samplings. In this study, several genetic variation differences between

good responders and non-responders were indeed noticed, in particular for *TPSD1* c.274G>A (p.Ala92Thr), *CPAMD8* c.1022G>A (p.Arg341Gln), and *OBP2A* c.367_370delCGCTinsTGCC (p.ArgTyr123CysHis). These differences were observed in tumour DNA of several non-responders but in none of the responders. After classification by Varsome, these three variants were characterised as those of uncertain significance. Of note, *TPSD1*, a gene encoding for a serine protease expressed in mast cells, is important for its optimal catalytic activity, and has been involved in autoimmune pathology (33, 34). *CPAMD8*, a gene encoding for the protease inhibitor I39 (alpha-2-macroglobulin), is necessary for innate and acquired immunity (35). Finally, *OBP2A* encodes for the odorant-binding protein 2a, a small extracellular protein important for transport of small hydrophobic molecules, such as steroids and lipids (36).

Several studies focusing on one or a few genes (such as *APC*, *KRAS* and *TP53*) confirmed the predictive value of cfDNA levels in therapeutic response monitoring in metastatic colorectal cancer, and the prognostic significance of post-operative cfDNA in early and locally advanced colon and rectal cancers (37–40). cfDNA has also been used to guide therapeutic decision-making in CC (37).

The genetic variations identified in *CPAMD8* and *OBP2A* in tumour DNA were also observed in cfDNA samples isolated from good responders. On the other hand, all good responders presented an insertion classified as VUS in the *ZXDB* gene (c.364_368delGAGGAinsAAGGC; p.GluGlu122LysAla) that was absent in all the non-responders. Interestingly, *ZXDB* encodes for the zinc finger X-linked protein which promotes the major histocompatibility complex (MHC) class I and II gene transcription (41).

The resistance mechanisms might be identified by obtaining individual resistant tumour lesions for molecular analysis. However, tumour tissue profiling has several limitations for the dynamic monitoring of disease progression and response to therapy: spatial and temporal tumour heterogeneity, difficulties to obtain repeated tissue samples.

Liquid biopsy has been proposed as a way to overcome these restrictions of tumour tissue profiling since the material can be collected with minimal invasiveness and repeated over time. Serial cfDNA measurements could complement the information available from routine imaging method assessments in the evaluation of tumour volume and response to chemotherapeutic agents (42). If the response to a particular treatment could be reliably assessed earlier, for example by serial cfDNA sampling, there could be an earlier switch to alternative therapy and a reduction of adverse effects resulting from inadequate therapy (37).

Using longitudinal plasma cfDNA analysis, we studied the evolution of the tumour genome in CC patients with the aim to identify mutations associated with the therapeutic response in good and non-responders. Interestingly, a pathogenic *APC* mutation (c.3944C>A; (p.Ser1315Ter) was detected on the third plasma sampling (a year after diagnosis and 6 months after the termination of therapy) of a non-responder. This particular mutation was not present at the time of diagnosis. The presence of a pathogenic *APC* mutation in plasma a year after surgery is in agreement with the assumption that this patient was classified as non-responder. This mutation could be a sign of an acquired treatment resistance, as *APC* is a known tumour suppressor gene that acts as an antagonist of the Wnt signalling pathway and is involved in cell migration and adhesion, transcriptional activation, and apoptosis.

Another non-responder presented a likely pathogenic *TP53* mutation (c.329G>C; p.Arg110Pro; rs11540654) in cfDNA in all its serial samplings and in tumour DNA. This outcome may represent

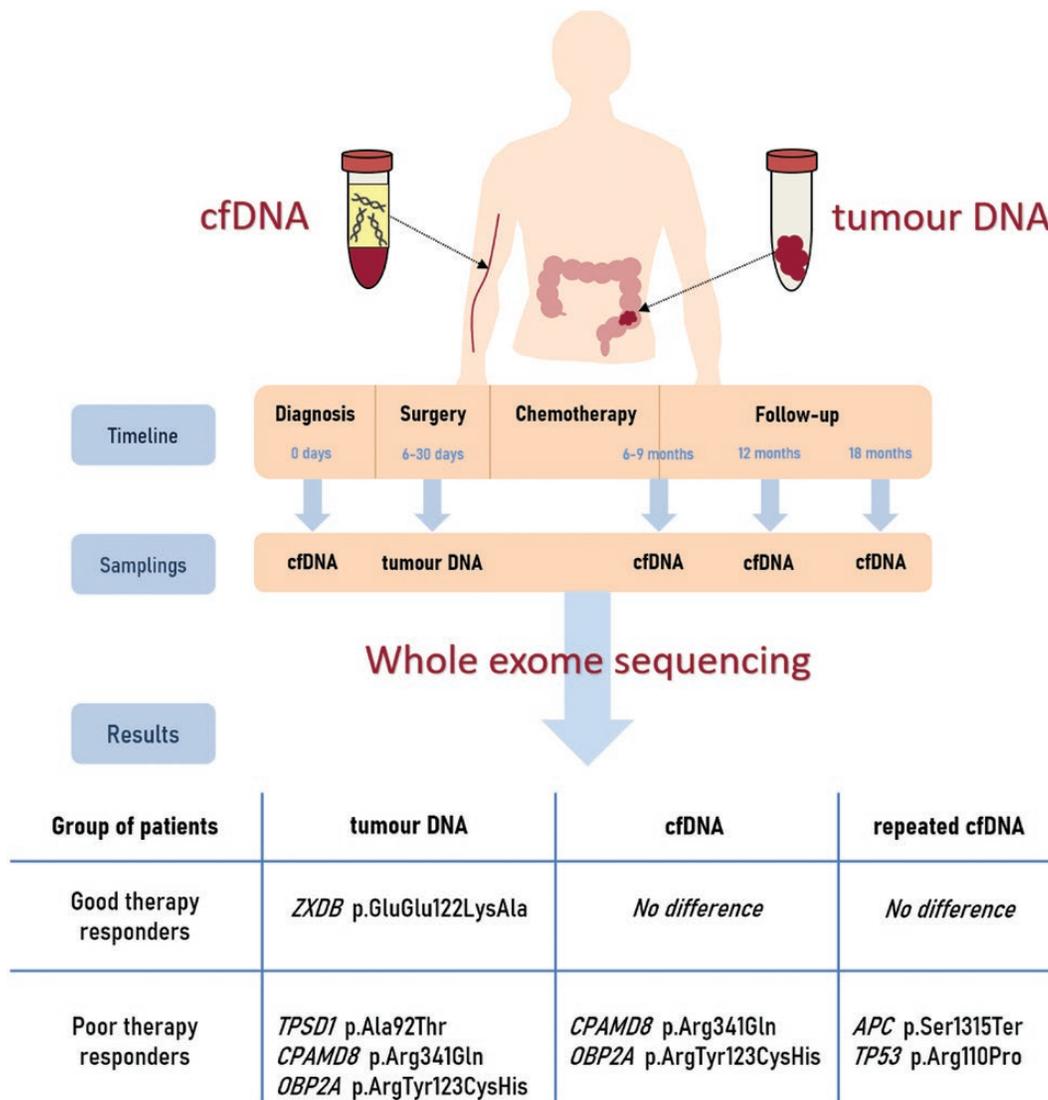


Figure 1. Workflow of the study.

another example of a poor response to therapy due to the persistence of a dysfunctional mutated tumour-derived p53 protein. These findings suggest that plasma cfDNA could be used as non-invasive approach for the genetic characterisation of CC patients and might be clinically useful for non-invasive tumour characterisation and therapeutic response monitoring.

Somatic mutations observed in cfDNA provide real-time data on the highly dynamic tumour evolution (43). Several authors have suggested that the analysis of post-operative cfDNA may identify patients who are at a higher risk of relapse (44–47). For example, cfDNA has been employed for the monitoring of secondary resistance to anti-epidermal growth factor receptor (anti-EGFR) treatment (8, 48–50). Similarly, several studies addressed that early cfDNA changes are associated with therapy responses in metastatic colorectal cancer patients (37, 51, 52). However, the selection of studied mutations was primarily based on mutations identified in tumour DNA at the time of diagnosis, which may constitute either an over or underestimation of the real patient situation due to the tumour heterogeneity/evolution as it was observed in this study. Monitoring the genetic changes in cfDNA during the therapy may improve the predictive accuracy of long-term prognosis, therefore the topic is worth further investigation.

Our hypothesis-generating study presents some limitations, the main being the limited number of patients included in the study. However, as it was recently shown in the study of Fleming et al. (53), even on relatively small population size, cfDNA may represent an important biomarker to identify those individuals that are at risk of early recurrence in CC. The outcomes hereby presented should be confirmed in a much larger study. Another limitation is rather low coverage depth. However, for WES this is the common approach and depth used. In addition, we analysed the read depth for each patient, sample (tumour or cfDNA), and for each identified gene to evaluate if for the variants highlighted and discussed in the study we had enough coverage. Although the overall read depth of samples is very low (especially for one tumour sample), the selected genes showed a high average read depth. For example, the median read depth for the *APC* gene in all samples (tumour tissue and cfDNA) was >1200. This higher read depth for selected genes has an impact on the capacity to identify variants in both tumour and plasma cfDNA. Another limitation is represented by the discrepancies of *APC*, *KRAS*, and *TP53* mutations between tumour DNA and plasma cfDNA at earlier stages. These differences could be partially explained by insufficient sequencing coverage in these genes. Moreover, it was possible to evaluate several time-points (sampling every 6 months) only in good responders. This was not always possible

for non-responders due to complications of the disease in these patients. A possible future approach could be to follow up non-responders even in shorter time points: this could support the observed results and consider added value. Serial cfDNA analysis should be included in future clinical trials in order to further assess this promising biomarker.

Conclusions

The identification of biomarkers associated with therapy response is crucial for the implementation of individualised treatment strategies in CC patients. Such biomarkers might stratify patients into groups with specific treatment and could finally promote the shift from a standardised to a personalised therapeutic approach. In this pilot study, we suggest that monitoring cfDNA over time may allow the identification of genetic variants closely associated with therapy response in CC patients.

Supplementary data

Supplementary data are available at *Mutagenesis* Online.

Funding

This project was supported by the Czech health research council of the Ministry of Health of the Czech Republic (AZV 17-30920A) (to K.C., A.N. and V.V.), and by the Grant Agency of Charles University (GAUK 302119 (to K.C.)).

Acknowledgements

We are thankful to the Charles University Research Centre program UNCE/MED/006 'University Centre of Clinical and Experimental Liver Surgery' provided by the Ministry of Education, Youth and Sports of the Czech Republic and by project no. CZ.02.1.01/0.0/0.0/16_019/0000787 'Fighting Infectious Diseases', awarded by the MEYS CR, financed from EFRR.

The work was partially supported by Lega Italiana per La Lotta contro I Tumori (LILT) (to B.P. and A.N.) and by the Oncobiome European H2020 research project (825410 (to A.N.)).

This article is based upon work from COST Action CA17118, www.transcoloncan.eu supported by COST (European Cooperation in Science and Technology). www.cost.eu

Author Contributions

Design study: V.V., B.P., A.N.; collection and preparation of samples: K.C., M.L., V.V., T.B.; performance of analysis: K.C., E.P., M.U., V.V., B.P.; data analysis: M.M.; interpretation of data: B.P., A.N., V.V.; original draft preparation: K.C., V.V., S.V., B.P., A.N.; writing – reviewing and editing: V.V., B.P., A.N., P.V., K.C., T.B., M.L., S.V., V.N. All authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement

T.B. received honoraria and research support from Roche, Bayer, Ipsen, Novartis, Merck, Bristol-Myers Squibb and Servier (unrelated to the present article). Other authors declare no conflicts of interest.

References

- Carethers, J. M. (2018) Risk factors for colon location of cancer. *Transl. Gastroenterol. Hepatol.*, 3, 76.
- Siegel, R. L., Miller, K. D. and Jemal, A. (2020) Cancer statistics, 2020. *CA Cancer J. Clin.*, 70, 7–30.
- Siegel, R. L., et al. (2020) Colorectal cancer statistics, 2020. *CA Cancer J. Clin.*, 70, 145–164.
- Goodwin, R. A. and Asmis, T. R. (2009) Overview of systemic therapy for colorectal cancer. *Clin. Colon Rectal Surg.*, 22, 251–256.
- Vodenkova, S., Buchler, T., Cervena, K., Veskrnova, V., Vodicka, P. and Vymetalkova, V. (2020) 5-Fluorouracil and other fluoropyrimidines in colorectal cancer: past, present and future. *Pharmacol. Ther.*, 206, 107447.
- El-Deiry, W. S., Goldberg, R. M., Lenz, H. J., et al. (2019) The current state of molecular testing in the treatment of patients with solid tumors, 2019. *CA Cancer J. Clin.*, 69, 305–343.
- Han, X., Wang, J. and Sun, Y. (2017) Circulating tumor DNA as biomarkers for cancer detection. *Genomics Proteomics Bioinformatics*, 15, 59–72.
- Zhou, Q., Perakis, S. O., Ulz, P., et al. (2020) Cell-free DNA analysis reveals POLR1D-mediated resistance to bevacizumab in colorectal cancer. *Genome Med.*, 12, 20.
- Li, G., Pavlick, D., Chung, J. H., et al. (2019) Genomic profiling of cell-free circulating tumor DNA in patients with colorectal cancer and its fidelity to the genomics of the tumor biopsy. *J. Gastrointest. Oncol.*, 10, 831–840.
- Crowley, E., Di Nicolantonio, F., Loupakis, F. and Bardelli, A. (2013) Liquid biopsy: monitoring cancer-genetics in the blood. *Nat. Rev. Clin. Oncol.*, 10, 472–484.
- 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.
- Jovelet, C., Ileana, E., Le Deley, M. C., et al. (2016) Circulating cell-free tumor DNA analysis of 50 genes by next-generation sequencing in the prospective MOSCATO Trial. *Clin. Cancer Res.*, 22, 2960–2968.
- Aravanis, A. M., Lee, M. and Klausner, R. D. (2017) Next-generation sequencing of circulating tumor DNA for early cancer detection. *Cell*, 168, 571–574.
- Cervena, K., Vodicka, P. and Vymetalkova, V. (2019) Diagnostic and prognostic impact of cell-free DNA in human cancers: systematic review. *Mutat. Res.*, 781, 100–129.
- Wu, Y. L., Sequist, L. V., Hu, C. P., et al. (2017) EGFR mutation detection in circulating cell-free DNA of lung adenocarcinoma patients: analysis of LUX-Lung 3 and 6. *Br. J. Cancer*, 116, 175–185.
- Parsons, H. A., Rhoades, J., Reed, S. C., et al. (2020) Sensitive detection of minimal residual disease in patients treated for early-stage breast cancer. *Clin. Cancer Res.*, 26, 2556–2564.
- Ling, Z. Q., Lv, P., Lu, X. X., et al. (2013) Circulating Methylated XAF1 DNA indicates poor prognosis for gastric cancer. *PLoS One*, 8, e67195.
- Cohen, J. D., Javed, A. A., Thoburn, C., et al. (2017) Combined circulating tumor DNA and protein biomarker-based liquid biopsy for the earlier detection of pancreatic cancers. *Proc. Natl. Acad. Sci. U. S. A.*, 114, 10202–10207.
- Abbosh, C., Birkbak, N. J., Wilson, G. A., et al.; TRACERx consortium; PEACE consortium. (2017) Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature*, 545, 446–451.
- Kim, S. T., Lira, M., Deng, S., et al. (2015) PIK3CA mutation detection in metastatic biliary cancer using cell-free DNA. *Oncotarget*, 6, 40026–40035.
- Osumi, H., Shinozaki, E., Takeda, Y., et al. (2019) Clinical relevance of circulating tumor DNA assessed through deep sequencing in patients with metastatic colorectal cancer. *Cancer Med.*, 8, 408–417.
- Vymetalkova, V., et al. (2018) Circulating cell-free DNA and colorectal cancer: a systematic review. *Int. J. Mol. Sci.*, 19, 3356.
- Marcuello, M., Vymetalkova, V., Neves, R. P. L., et al. (2019) Circulating biomarkers for early detection and clinical management of colorectal cancer. *Mol. Aspects Med.*, 69, 107–122.
- Palmirotta, R., Lovero, D., Cafforio, P., Felici, C., Mannavola, F., Pellè, E., Quaresmini, D., Tucci, M. and Silvestris, F. (2018) Liquid biopsy of cancer: a multimodal diagnostic tool in clinical oncology. *Ther. Adv. Med. Oncol.*, 10, 1758835918794630.
- Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Albarca Aguilera, M., Meyer, R. and Massouras, A. (2019) VarSome: the human genomic variant search engine. *Bioinformatics*, 35, 1978–1980.
- Jia, N., Sun, Z., Gao, X., et al. (2019) Serial monitoring of circulating tumor DNA in patients with metastatic colorectal cancer to predict the therapeutic response. *Front. Genet.*, 10, 470.
- Lin, J. K., Lin, P. C., Lin, C. H., Jiang, J. K., Yang, S. H., Liang, W. Y., Chen, W. S. and Chang, S. C. (2014) Clinical relevance of alterations in quantity and quality of plasma DNA in colorectal cancer patients: based

- on the mutation spectra detected in primary tumors. *Ann. Surg. Oncol.*, 21(Suppl 4), S680–S686.
28. Sylvester, B. E. and Vakiani, E. (2015) Tumor evolution and intratumor heterogeneity in colorectal carcinoma: insights from comparative genomic profiling of primary tumors and matched metastases. *J. Gastrointest. Oncol.*, 6, 668–675.
 29. He, Y., Ma, X., Chen, K., Liu, F., Cai, S., Han-Zhang, H., Hou, T., Xiang, J. and Peng, J. (2020) Perioperative circulating tumor DNA in colorectal liver metastases: concordance with metastatic tissue and predictive value for tumor burden and prognosis. *Cancer Manag. Res.*, 12, 1621–1630.
 30. Sottoriva, A., Kang, H., Ma, Z., *et al.* (2015) A big bang model of human colorectal tumor growth. *Nat. Genet.*, 47, 209–216.
 31. Blank, A., Roberts, D. E. 2nd, Dawson, H., Zlobec, I. and Lugli, A. (2018) Tumor heterogeneity in primary colorectal cancer and corresponding metastases. does the apple fall far from the tree? *Front. Med. (Lausanne)*, 5, 234.
 32. Wood, L. D., Parsons, D. W., Jones, S., *et al.* (2007) The genomic landscapes of human breast and colorectal cancers. *Science*, 318, 1108–1113.
 33. Pallaoro, M., Fejzo, M. S., Shayesteh, L., Blount, J. L. and Caughey, G. H. (1999) Characterization of genes encoding known and novel human mast cell tryptases on chromosome 16p13.3. *J. Biol. Chem.*, 274, 3355–3362.
 34. Caughey, G. H. (2007) Mast cell tryptases and chymases in inflammation and host defense. *Immunol. Rev.*, 217, 141–154.
 35. Cheong, S. S., Hentschel, L., Davidson, A. E., *et al.* (2016) Mutations in CPAMD8 cause a unique form of autosomal–recessive anterior segment dysgenesis. *Am. J. Hum. Genet.*, 99, 1338–1352.
 36. Tcatchoff, L., Nespoulous, C., Pernollet, J. C. and Briand, L. (2006) A single lysyl residue defines the binding specificity of a human odorant-binding protein for aldehydes. *FEBS Lett.*, 580, 2102–2108.
 37. Tie, J., Kinde, I., Wang, Y., *et al.* (2015) Circulating tumor DNA as an early marker of therapeutic response in patients with metastatic colorectal cancer. *Ann. Oncol.*, 26, 1715–1722.
 38. Wang, Y., *et al.* (2019) Prognostic potential of circulating tumor DNA measurement in postoperative surveillance of nonmetastatic colorectal cancer. *JAMA Oncol.*
 39. Tie, J., Cohen, J. D., Wang, Y., *et al.* (2019) Serial circulating tumour DNA analysis during multimodality treatment of locally advanced rectal cancer: a prospective biomarker study. *Gut*, 68, 663–671.
 40. Tie, J., Cohen, J. D., Wang, Y., *et al.* (2019) Circulating tumor DNA analyses as markers of recurrence risk and benefit of adjuvant therapy for stage III colon cancer. *JAMA Oncol.*, 5, 1710–1717.
 41. Al-Kandari, W., Koneni, R., Navalgund, V., Aleksandrova, A., Jambunathan, S. and Fontes, J. D. (2007) The zinc finger proteins ZXDA and ZXDC form a complex that binds CIITA and regulates MHC II gene transcription. *J. Mol. Biol.*, 369, 1175–1187.
 42. Bi, F., Wang, Q., Dong, Q., Wang, Y., Zhang, L. and Zhang, J. (2020) Circulating tumor DNA in colorectal cancer: opportunities and challenges. *Am. J. Transl. Res.*, 12, 1044–1055.
 43. Diehl, F., Schmidt, K., Choti, M. A., *et al.* (2008) Circulating mutant DNA to assess tumor dynamics. *Nat. Med.*, 14, 985–990.
 44. Kidess, E., Heirich, K., Wiggin, M., *et al.* (2015) Mutation profiling of tumor DNA from plasma and tumor tissue of colorectal cancer patients with a novel, high-sensitivity multiplexed mutation detection platform. *Oncotarget*, 6, 2549–2561.
 45. Reinert, T., Schöler, L. V., Thomsen, R., *et al.* (2016) Analysis of circulating tumour DNA to monitor disease burden following colorectal cancer surgery. *Gut*, 65, 625–634.
 46. Zhou, J., Chang, L., Guan, Y., Yang, L., Xia, X., Cui, L., Yi, X. and Lin, G. (2016) Application of circulating tumor DNA as a non-invasive tool for monitoring the progression of colorectal cancer. *PLoS One*, 11, e0159708.
 47. Ng, S. B., Chua, C., Ng, M., *et al.* (2017) Individualised multiplexed circulating tumour DNA assays for monitoring of tumour presence in patients after colorectal cancer surgery. *Sci. Rep.*, 7, 40737.
 48. Morelli, M. P., Overman, M. J., Dasari, A., *et al.* (2015) Characterizing the patterns of clonal selection in circulating tumor DNA from patients with colorectal cancer refractory to anti-EGFR treatment. *Ann. Oncol.*, 26, 731–736.
 49. Siravegna, G., Mussolin, B., Buscarino, M., *et al.* (2015) Clonal evolution and resistance to EGFR blockade in the blood of colorectal cancer patients. *Nat. Med.*, 21, 795–801.
 50. Takegawa, N., Yonesaka, K., Sakai, K., *et al.* (2016) HER2 genomic amplification in circulating tumor DNA from patients with cetuximab-resistant colorectal cancer. *Oncotarget*, 7, 3453–3460.
 51. Garlan, F., Laurent-Puig, P., Sefrioui, D., *et al.* (2017) Early evaluation of circulating tumor DNA as marker of therapeutic efficacy in metastatic colorectal cancer patients (PLACOL Study). *Clin. Cancer Res.*, 23, 5416–5425.
 52. Thomsen, C. B., Hansen, T. F., Andersen, R. F., Lindebjerg, J., Jensen, L. H. and Jakobsen, A. (2018) Monitoring the effect of first line treatment in RAS/RAF mutated metastatic colorectal cancer by serial analysis of tumor specific DNA in plasma. *J. Exp. Clin. Cancer Res.*, 37, 55.
 53. Fleming, C. A., O’Leary, D. P., Wang, J. and Redmond, H. P. (2020) Association of observed perioperative cell-free DNA dynamics with early recurrence in patients with colon cancer. *JAMA Surg.*, 155, 168–170.