ORIGINAL ARTICLE

Analysis of circulating tumour DNA to monitor disease burden following colorectal cancer surgery

Thomas Reinert,[1] Lone V Schøler,[1] Rune Thomsen,[1] Heidi Tobiasen,[1] Søren Vang,[1] Iver Nordentoft,[1] Philippe Lamy,[1] Anne-Sofie Kannerup,[2] Frank V Mortensen,[2] Katrine Stribolt,[3] Stephen Hamilton-Dutoit,[3] Hans J Nielsen,[4] Søren Laurberg,[5] Niels Pallisgaard,[6] Jakob S Pedersen,[1] Torben F Ørntoft,[1] Claus L Andersen[1]

To cite: Reinert T, Schøler LV, Thomsen R, et al. Gut 2016;**65**: 625–634.

ABSTRACT

Objective To develop an affordable and robust pipeline for selection of patient-specific somatic structural variants (SSVs) being informative about radicality of the primary resection, response to adjuvant therapy, incipient recurrence and response to treatment performed in relation to diagnosis of recurrence.

Design We have established efficient procedures for identification of SSVs by next-generation sequencing and subsequent quantification of 3–6 SSVs in plasma. The consequence of intratumour heterogeneity on our approach was assessed. The level of circulating tumour DNA (ctDNA) was quantified in 151 serial plasma samples from six relapsing and five non-relapsing colorectal cancer (CRC) patients by droplet digital PCR, and correlated to clinical findings.

Results Up to six personalised assays were designed for each patient. Our approach enabled efficient temporal assessment of disease status, response to surgical and oncological intervention, and early detection of incipient recurrence. Our approach provided 2–15 (mean 10) months' lead time on detection of metastatic recurrence compared to conventional follow-up. The sensitivity and specificity of the SSVs in terms of detecting postsurgery relapse were 100%.

Conclusions We show that assessment of ctDNA is a non-invasive, exquisitely specific and highly sensitive approach for monitoring disease load, which has the potential to provide clinically relevant lead times compared with conventional methods. Furthermore, we provide a low-coverage protocol optimised for identifying SSVs with excellent correlation between SSVs identified in tumours and matched metastases. Application of ctDNA analysis has the potential to change clinical practice in the management of CRC.

Significance of this study

What is already known on this subject?
► Solid cancers, including colorectal cancers (CRC), release DNA into circulation.
► CRCs harbour somatic genetic alterations, including both single-base substitutions and larger somatic structural variations (SSVs).
► Early detection of recurrence and subsequent intervention has been shown to be associated with increased patient survival.

What are the new findings?
► We are the first to report a low-coverage, robust and yet very efficient pipeline for identification and prioritising of patient-specific SSVs according to their potential of being informative in monitoring patients with CRC following initial surgery.
► We report a procedure, which includes controls for cell-free DNA (cfDNA) purification efficiency, leucocyte DNA contamination and cfDNA quantity. Importantly, these controls enable the lower circulating tumour DNA (ctDNA) detection limit for each plasma sample to be assessed prior to analysis.
► We investigate both relapsing and non-relapsing patients and demonstrate excellent sensitivity and specificity of ctDNA analysis in detecting relapse. Relapse is on average detected 10 months earlier than by conventional follow-up.
► We report the ability of ctDNA monitoring to inform about radicality of the primary resection, response to adjuvant therapy and response to interventions performed in relation to diagnosis of recurrence. We compare the performance of ctDNA with the performance of cfDNA and carcinoembryonic antigen from the same time points.

INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer worldwide.[1] Approximately two-thirds of patients will present with potentially curable disease (by surgery±adjuvant therapies),[2] but in spite of curatively intended treatment 30–40% of these patients will experience recurrence of disease.[3] Surveillance for recurrence of CRC after curatively intended surgery is clinically important because early detection of recurrence and subsequent intervention has been shown to be associated with increased patient survival.[4 5] Current surveillance guidelines recommend a combination of different tools for effective surveillance. These generally include clinical assessment, serum carcinoembryonic antigen (CEA) testing, colonoscopy and CT.[6–8]

Non-invasive analysis of circulating tumour DNA is an emerging tool that has the potential to improve the field of postsurgery surveillance. It is

Significance of this study

How might it impact on clinical practice in the foreseeable future?

► Efficient monitoring of ctDNA has the potential become a practice-changing tool. It has the potential to create a critical window of opportunity for intervention at time points where curative modalities are still an option. Furthermore, it may provide sensitive and timely information on patient-specific response to surgical and oncological interventions. Consequently, in the future it may become an efficient tool for personalising and optimising the postsurgery clinical management of patients with cancer.

based on the well-recognised observation that solid tumours, including CRC, release DNA fragments into plasma.[9–11] Recent sequencing studies have shown that virtually all CRCs harbour somatic genetic alterations, including both single-base substitutions and larger somatic structural variations (SSVs).[12 13] Importantly, these somatic mutations occur at negligible frequencies in normal cell populations and therefore have the potential to be used as exquisitely specific biomarkers for detection and quantification of circulating tumour DNA (ctDNA).[14] Several recent reports have demonstrated the feasibility of detecting ctDNA in early-stage and late-stage malignancies, and for monitoring resistance to therapy in the metastatic setting.[15–17] A challenge for ctDNA analysis is the identification of the somatic mutations to be used as markers. Typically either point mutations in hotspot genes or patient-specific SSVs have been used. The challenge of hotspot mutations is the specificity of the assays and that only a fraction of patients are mutated. SSVs, by contrast, are found in all patients, and specificity is generally no problem due to the major genomic change imposed by structural alterations. The challenge of SSVs is,

however, that they are patient specific and novel SSVs have to be identified for each patient.

The aims of the present study were to establish an affordable, robust and yet very efficient pipeline for the identification of patient-specific SSVs, to devise an algorithm for prioritising SSVs according to their potential of being informative on recurrence, to devise a pipeline for designing quantitative assays to SSVs, to devise a quality control (QC) procedure enabling the sensitivity of each ctDNA analysis to be assessed and to investigate the impact of intratumour heterogeneity on the whole setup. To evaluate whether our approach enabled earlier detection of impending recurrence than standard clinical follow-up programmes, we applied it to a set of 11 patients with CRC, who all had been resected with curative intent. Nevertheless, six of the patients experienced disease recurrence. We also evaluated the ability of ctDNA monitoring to inform about radicality of the primary resection, response to adjuvant therapy and response to interventions performed in relation to diagnosis of recurrence. Finally, the performance of ctDNA was compared with CEA at the same time points.

SUBJECTS AND METHODS

From 2005 to 2007, tumour tissue was consecutively collected from 118 patients at the Department of Surgery, Aarhus University Hospital. Blood samples were collected at day 0 (preoperative), 8, 30 and every three months until death, patient withdrawal from the study or month 36, whichever came first. The last blood sample was collected in 2010. From this cohort, six relapsing and five non-relapsing patients were retrospectively selected for this study. These 11 patients contributed 151 blood samples. From one of the relapsing patients, metastatic tissue was accessible. An additional three patients with matched tumour and metastatic tissue were included in the study. Tumour and metastatic tissue samples were obtained fresh from surgery, embedded in Tissue-Tek (Sakura Prohosp), and immediately snap frozen in liquid nitrogen. Matched germline DNA samples were obtained from peripheral blood leucocytes.

Table 1 Demographic and clinical characteristics of colorectal cancer patients

Patient	Gender	Age	Localisation	Stage	TNMV	Initial surgery/RFA	Recurrence diagnosed (month postoperative)	Treatment (month postoperative) Chemotherapy	Treatment (month postoperative) Surgery/RFA/radiation
1	Male	81	Colon	II	T3N0M0V0	Colon	None	No	None
2	Male	77	Rectum	II	T3N0M0V0	Rectum	None	No	None
4	Female	81	Colon	III	T3N2M0V1	Colon	Local (15), liver and lung (25)	Adjuvant (1)	Local recurrence (17)*
5	Male	76	Colon	II	T3N0M0V0	Colon	None	No	None
6	Male	70	Rectum	II	T3N0M0V1	Rectum	None	No	None
8	Male	59	Rectum	III	T3N2M0V0	Rectum	Lung (28), liver (39)	No	Lobectomy (30)/lung RFA (31, 36)
10	Female	73	Rectum	IV	T3N1M1V1	Rectum/liver	Liver (32)	Chemo (5)	Liver RFA (33)
15	Male	65	Rectum	II	T2N0M0V0	Rectum	Local (24)	No	Local recurrence (26)
16	Female	56	Rectum	I	T2N0M0V0	Rectum	Liver (19)	Chemo (20)	None
18	Female	63	Colon	IV	T3N0M1V0	Colon and liver	Liver (35)	No	Partial hepatectomy (38)
19	Male	59	Colon	IV	T3N1M1V1	Colon/liver	None	Chemo (4)	None
24†	Female	61	Colon	II	T3N0M0V0	Colon	Liver (7)	No	Partial hepatectomy (9)/RFA (9, 23)
28†	Female	47	Rectum	III	T4N2M0V0	Rectum	Liver (29)	Chemo (32)	Partial hepatectomy (31)
29†	Female	60	Colon	IV	T4N0M1V0	Colon and liver /liver	None	Chemo (1, 5)	None

*Supravaginal hysterectomy, small bowel resection and salpingo-oophorectomy.
†No plasma samples available.
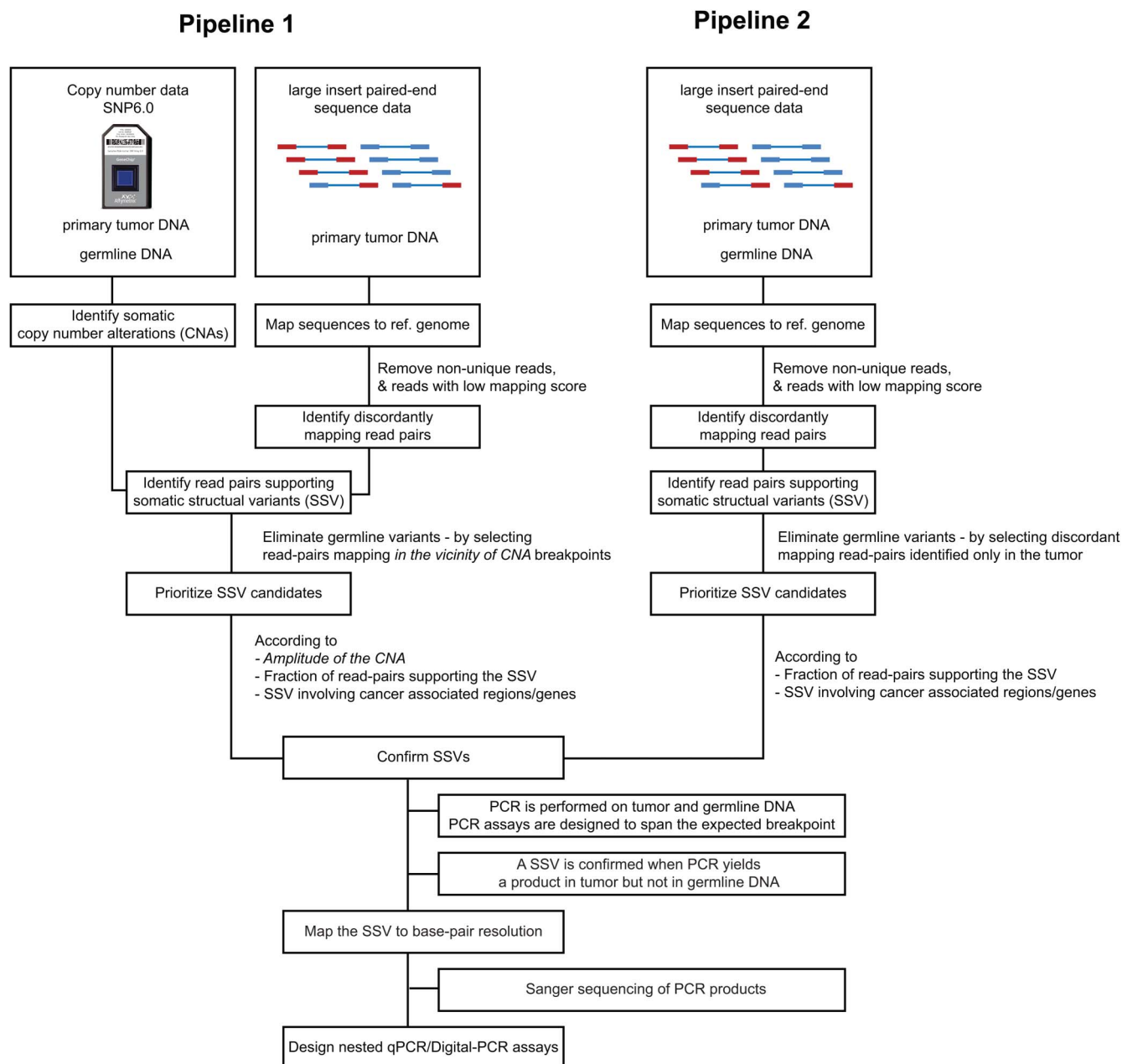RFA, radiofrequency ablation.

Pipeline 1

Pipeline 2

Figure 1 Pipelines for identification of somatic structural variations (SSVs). Two different pipelines for identification of SSVs in the primary tumour were applied. Both pipelines were performed without knowledge of patient outcome. In pipeline 1, only the tumour was mate-pair sequenced and used for structural variant calling. To pinpoint somatic variants, we superimposed somatic copy number alteration data, derived from single-nucleotide polymorphism (SNP) array analysis of paired tumour and germline samples, onto the sequence data. This enabled identification of discordantly mapping read pairs consistent with somatic copy number gains and losses. This approach typically mapped the associated genomic breakpoints to narrow regions ranging in size from 500 to 3000 bases. Subsequently, the breakpoint was mapped to base-pair resolution by Sanger sequencing. Pipeline 2 differed by using another method for identification of somatic structural variants. Mate-pair data were generated from both tumour and germline DNA, thereby enabling identification of somatic structural variants directly from the sequence data; hence, pipeline 2, in contrast to pipeline 1, allowed identification of additional types of variants in addition to those associated with deletions and amplifications, for example, inversions and balanced intra-chromosomal and inter-chromosomal translocations. The precise genomic position of the breakpoints was mapped as in pipeline 1.

Formalin fixed paraffin embedded (FFPE) tumour and metastatic tissues were obtained from the Institute of Pathology at Aarhus University Hospital. Tissue samples had a median tumour percentage of 75 (range 50–90). Cancer content was assessed by H&E evaluation of sections cut before and after those used for extraction (see online supplementary table S1). Information on postsurgery clinical intervention was available on all patients (table 1).

Identification of somatic copy number alterations by single-nucleotide polymorphism-array data

Matched germline and tumour DNA were profiled using single-nucleotide polymorphism (SNP) V.6.0 arrays (Affymetrix). Initial QC of the arrays was performed using the Birdsuite software.[18] Tumour-specific copy number alterations (CNAs) were derived from each tumour/normal pair (see online supplementary methods).

Table 2 SSVs identified by pipelines 1 and 2

	Non-recurrence						Recurrence			
	Patient 1	Patient 2	Patient 5	Patient 6	Patient 19	Patient 8	Patient 10	Patient 15	Patient 16	Patient 18
Pipeline 1	13	25	0	13	7	8	7	12	4	9
Pipeline 2	ND	ND	5	ND	ND	13	14	ND	14	19
Unique SSVs*	13	25	5	13	7	14	15	12	16	20
Number of SSVs analysed in plasma†	5	5	4	4	6	5	4	2	3	6
Number of SSVs detected in plasma	0	2‡	0	3‡	3‡	3	3	2	3	3

*The total number of unique SSVs identified (the union of pipelines 1 and 2).
†For each patient, 4–8 of the unique SSVs (those supported by most reads or affecting loci harbouring genes known to drive CRC development, and preferably mapping to different chromosomes) were selected. Their breakpoints were mapped to nucleotide level by Sanger sequencing, and subsequently, ddPCR assays were designed to the SSVs with the breakpoints mapping to non-repetitive sequences. Only the SSVs for which high-performance ddPCR assays could be produced were analysed in plasma.
‡SSV was only detected in the plasma sample drawn prior to surgery.
CRC, colorectal cancer; ddPCR, droplet digital PCR; ND, not done; SSVs, somatic structural variations.

Illumina mate pair libraries and next-generation sequencing

Whole-genome sequencing was done using Illumina mate pair and TruSeq DNA Sample Preparation Kit. Illumina sequencing was performed on the Illumina HiSeq2000 platform (see online supplementary methods). On average, 28.5 million read pairs were sequenced per sample yielding sequence and physical read depths of 1.6 and 27.0, respectively. Detailed information for each sample is listed in online supplementary table S2.

Analysis of Illumina sequence data

To identify candidate structural variants, the software tool BreakDancer was applied to the final alignments.[19] Bed-files were produced to visualise the structural variants using the Integrative Genomics Viewer browser.[20] To confirm the common origin of matched germline and tumour samples, mitochondrial SNPs were called from all sequence libraries and used for hierarchical clustering (see online supplementary methods).

Isolation and quantification of DNA

DNA was extracted from fresh frozen tissue using the Puregene DNA purification kit (Gentra Systems), from FFPE tissue using the QiaAMP DNA FFPE Tissue kit (56404), from 2 to 4 mL plasma by QIAamp DNA Blood Midi Kit (51185) or QIAamp Circulating Nucleic Acid Kit (55114) (see online supplementary methods). As internal control for DNA purification efficiency, we spiked in a fixed number of copies of a PCR fragment from the DNA binding protein CPP1, expressed specifically in soybean, to each lysed plasma sample (see online supplementary table S3).[21] The number of CPP1, cell-free DNA (cfDNA) and ctDNA template copies in each purified sample was quantified by droplet digital PCR using assays specific to CPP1, two reference regions on chromosomes 3 and 7 that only rarely show CNAs in CRC and tumour genomic alterations (see online supplementary table S3).[22] Lymphocyte DNA contamination was estimated by an assay targeting the VDJ rearranged IGH locus specific for B cells (see online supplementary table S3).[21] Together these measurements were used as QCs of the plasma DNA and to assess the ctDNA detection limit for each sample, that is, 1/(cfDNA quantity per sample). Also, 2 out of 151 plasma samples showed minor contamination with lymphocyte DNA. Their cfDNA levels did not deviate from the rest, making us to flag, rather than exclude, them (see online supplementary figure S1).

Validation and Sanger sequencing

SSVs identified by mate-pair sequencing were confirmed by a breakpoint spanning PCR of tumour and matched germline DNA using the TEMPase PCR Kit (Ampliqon) according to the manufacturer's instructions. The patient-specific primers are listed in online supplementary table S3. All breakpoints were subsequently mapped to nucleotide level using Sanger sequencing on a 3130x Genetic Analyzer (Applied Biosystem).

Amplification of SSVs by multiplex nested PCR

For the analysis of patients 10 and 16, 12 cycles of nested PCR were carried out using 90% of the DNA purified from the plasma samples (see online supplementary methods).

DNA quantification by droplet digital PCR

DNA samples were analysed on a QX100 droplet digital PCR (ddPCR) system according to the manufacturer's instructions (Bio-Rad, Pleasanton, California, USA). Linearity and sensitivity of the patient-specific assays for detection of genomic SSVs were assessed using a six-point dilution series of tumour DNA (4000, 1000, 250, 62.5, 15.6 and 3.9 genome equivalents (GEs)) in a constant pool of 20 000 GEs of matched germline DNA. When analysing the clinical plasma samples, ∼20% of the DNA isolated from 2 to 4 mL plasma was used as template for each patient-specific assay. For the samples tested with the nested approach (patients 10 and 16), 2 µL diluted nested PCR product was used as template. FFPE tissues were analysed by ddPCR using 40 ng of DNA. Primer and probe sequences are listed in online supplementary table S3.

CEA analysis

The CEA analyses were performed on a Cobas e601 (Roche), according to the manufacturer's recommendation using 500 mL serum, at the Department of Clinical Biochemistry, Aarhus University Hospital. The threshold levels were set to 4.0 and 6.0 µg/L for non-smokers and smokers, respectively.

RESULTS
Identification of SSVs

We applied two different pipelines to identify SSVs (figure 1). Pipeline 1 was based on a combination of tumour DNA large-insert paired-end sequencing and paired tumour/germline SNP array-derived copy number analysis. In pipeline 2, SSVs were identified by large-insert paired-end sequencing of matched tumour and germline DNA. The pipelines were run completely independent of patient outcome and are well suited for prospective follow-up. Pipeline 1 was applied to tumour DNA from all patients from which serial postsurgery plasma samples were available and pipeline 2 to five of these. In all five patients,
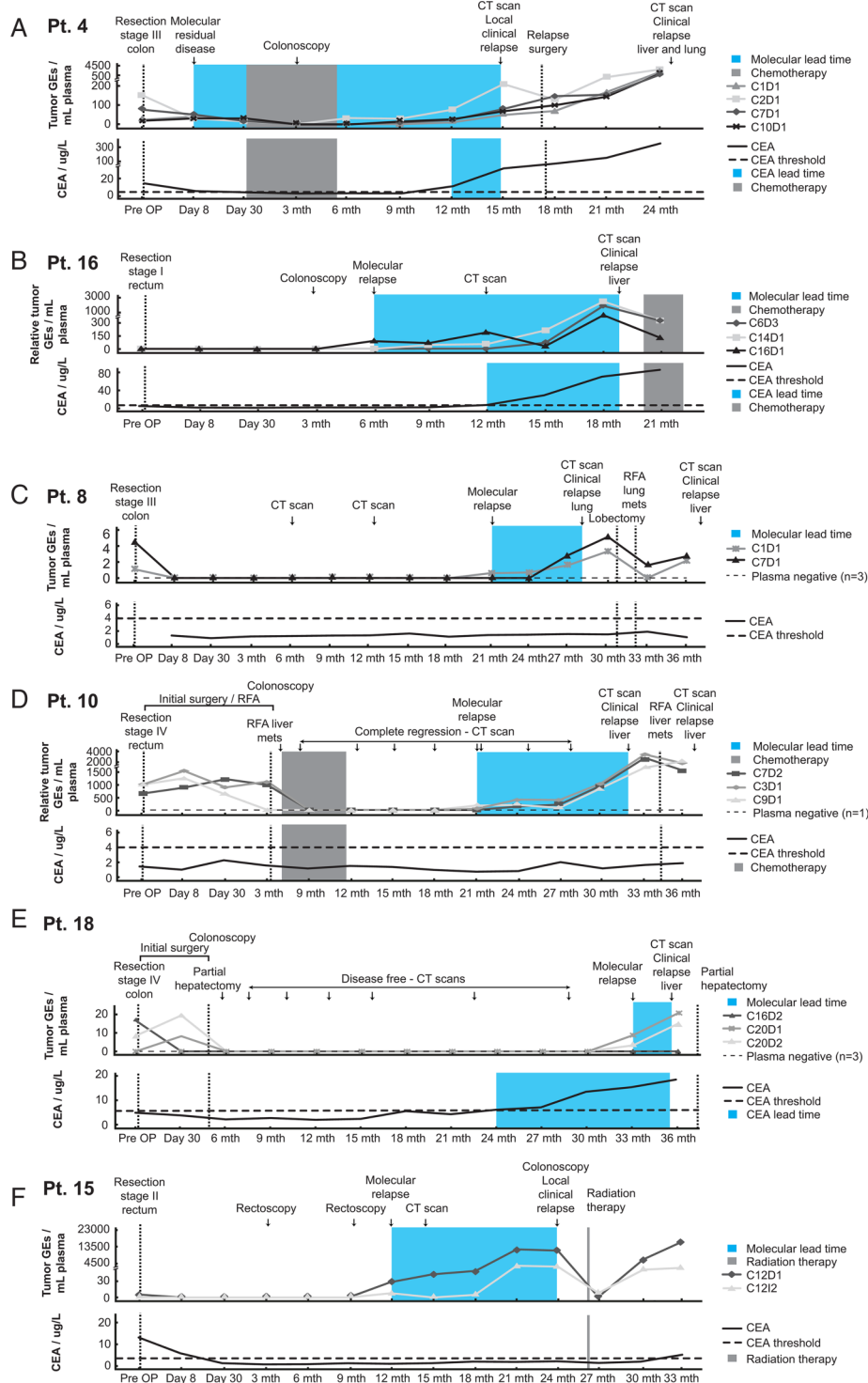
Figure 2 Clinical application of circulating tumour DNA (ctDNA) and carcinoembryonic antigen (CEA) for monitoring colorectal cancer in patients with residual disease and/or recurrence. Cell-free DNA (cfDNA) and CEA were isolated from serial plasma and serum samples, respectively. Samples were collected every third month starting prior to surgery and ending at month 36 postsurgery or in relation to recurrence of disease. TaqMan assays, amplifying patient-specific somatic structural variations identified by mate-pair sequencing were designed and applied to cfDNA to enable quantification of the level of ctDNA. Shown are the results of monitoring the level of ctDNA (upper figure) and CEA (lower figure) in patients surgically treated with curative intend for (A) a stage III colon cancer (Pt. 4); (B) a stage I rectum cancer (Pt. 16); (C) a stage III rectum cancer (Pt. 8); (D) a stage IV rectum cancer with focal metastasis in the lung. The initial treatment included resection of the primary tumour and radiofrequency ablation (RFA) to eliminate the lung metastasis (Pt. 10); (E) a stage IV colon cancer with focal metastasis in the liver. The initial treatment included a colon resection to remove the primary tumour and a partial hepatectomy to eliminate the liver metastasis (Pt. 18); (F) a stage II rectum cancer with subsequent local recurrence treated with radiation therapy (Pt. 15). The quantified levels of ctDNA are plotted as tumour genome equivalents (GEs). Relative tumour GEs are from patients where nested multiplex PCR were used. Data are only shown for informative assays (see online supplementary table S4 for a complete list of assays). Vertical dotted lines indicate surgery or RFA. Grey shaded regions indicate chemotherapy. Arrows indicate radiological and molecular examinations, and they were negative unless specified otherwise. Blue shaded regions indicate lead time. CEA threshold levels are indicated by horizontal dotted lines. Threshold values are 4 and 6 µg/L for non-smokers and smokers, respectively.

pipeline 2 identified more SSVs than pipeline 1 (table 2), primarily because pipeline 2 was not restricted to identifying SSVs causing CNAs, but also identified inversions and balanced intra-chromosomal and inter-chromosomal translocations.

The overlap of SSVs between pipeline 1 and 2 ranged from 0 to 8. All SSVs identified by pipeline 1 and supported by more than one read pair were also identified by pipeline 2 (see online supplementary table S4). Taken together the pipelines detected five or more SSVs (average 15.5) per patient (table 2). For future studies, pipeline 2 is our preferred choice as it in practice provides more options for selecting informative SSVs.

When selecting SSVs to be tested as ctDNA biomarker candidates, we applied an algorithm prioritising SSVs supported by most read pairs and/or affecting loci harbouring genes known to drive CRC development. SSVs with these characteristics have a high probability of being present in the majority of cancer cells and are therefore likely to be present in later recurrences. Importantly, the SSV selection procedure was performed completely independent of patient outcome, that is, the exact same criteria were used for recurrence and non-recurrence patients. For each patient, 4–8 SSVs were selected and breakpoints were validated by Sanger sequencing. In order to facilitate ddPCR assay design and maximise specificity, only SSVs with breakpoints occurring in non-repetitive DNA sequences were taken on for assay design. For each patient, a selection of top-prioritised and validated SSVs was selected for ddPCR design (table 2 and online supplementary table S4).

Development of ddPCR biomarker assays from SSVs

ddPCR assays with amplicon lengths below 100 nucleotides were designed to PCR and Sanger sequence validated SSVs (see online supplementary tables S3 and S4). To simulate detection of ctDNA in plasma, where it constitutes only a minority of the total cfDNA, the sensitivity and the linearity of all ddPCR assays were thoroughly evaluated in a high background of germline DNA. All included assays consistently detected down to one GE of tumour DNA in a background of 20 000 GEs (=0.005%) of matched germline DNA. The linearity of the assays, measured by R^2, ranged from 0.9977 to 0.9998 across three orders of magnitude, showing that the ddPCR assays were highly sensitive, and had a linear range of quantification from 1 to 4000 GEs in large excess of germline DNA. The numbers of cfDNA templates per ddPCR reaction ranged from 42 to 11 583 GEs, with a median of 730 across the 151 plasma samples. Accordingly, the minimal fraction of ctDNA detectable in the clinical plasma samples ranged from 0.009% to 2.4%, median 0.13% (see online supplementary figures S1 and S2).

Tumour monitoring using ctDNA and cfDNA

Retrospectively, we studied 151 serial plasma samples collected during a 36-month follow-up period. They originated from 11 patients with CRC initially treated with curatively intended surgery for their stage I–IV disease (table 1). Six of the patients experienced relapse of disease and five did not. To detect and quantitate changes in the level of ctDNA during follow-up, we employed 2–6 tumour-specific ddPCR assays for each patient. The ctDNA analyses identified relapse in 6/6 relapsing and in 0/5 non-relapsing patients, yielding a sensitivity of 100% and a specificity of 100% (figures 2 and 3).

Furthermore, ctDNA was detected retrospectively in the blood samples drawn prior to surgery in all patients, except one stage 1 and two stage 2 (figures 2 and 3). Notably, the quantified level of mutant DNA in serial postsurgery plasma samples showed an intimate relationship with the clinical disease course.

The level generally decreased upon intervention, for example, resection, radiofrequency ablation and chemotherapy, at least for a short while, for then rapidly to increase in the months up to diagnosis of clinical recurrence.

Accordingly, for all six relapsing patients, the ctDNA analysis detected the incipient recurrence earlier than the conventional diagnostic modalities (eg, CT scanning). The average lead time was 10 months, ranging from 2 to 15 months (figure 2). Most importantly, no ctDNA was detected in any postsurgery plasma sample from the non-relapsing patients (figure 3). Notably, equal amounts of cfDNA were analysed for the relapsing and non-relapsing patients (see online supplementary figure S3A–C). Likewise, the number of SSVs investigated for the non-relapsing patients was the same or higher than for the relapsing patients (table 2).

In contrast to ctDNA, only a limited relationship between the quantified cfDNA levels and the clinical disease course was observed (see online supplementary figures S4 and S5). Due to high fluctuations in the quantified cfDNA levels during the follow-up period, but particularly within the first 30 days after surgery, it was for many patients difficult to set a cfDNA threshold for calling the relapse events unequivocally using cfDNA. Nevertheless, for 5/6 relapsing patients, an increase in the level of cfDNA was observed around the time when the relapse was clinically diagnosed. We observed a positive correlation between the cfDNA and ctDNA levels (Spearman r=0.37, p<0.0000), and even more so when only samples positive for ctDNA were included (Spearman r=0.67, p<0.0000) (see online supplementary figure S6A,B). As expected, the correlation increased as the level of ctDNA increased (see online supplementary figure S6C, D).

The different patient-specific tumour markers generally showed similar dynamic patterns in plasma (figures 2 and 3). However, in some cases, we also observed discordant patterns providing evidence of clonal heterogeneity and/or the existence of multiple lesions. The ctDNA analysis of patient 4 indicated that the initial surgery was not radical and while adjuvant chemotherapy initially suppressed all four mutant markers, they all came back up when therapy ended (figure 2A). Of importance though, not at the same pace, indicating the existence of more than one relapsing clone. Indeed, this was later confirmed as resection of a diagnosed local recurrence only caused a single ctDNA marker to decrease, while the others continued to increase. A subsequent CT scan identified metastases in the liver and lungs.

A limited number of our plasma samples were ctDNA marker negative, even though it was clear from the clinical disease course that disease existed. These were all from time points where the clinical tumour burden was very low, for example, prior to surgery of a patient diagnosed with a stage I or II tumour (figures 2B and 3A, C), or in a period postresection of the primary tumour (figure 2B–F). It is likely that ctDNA at these time points was too low to be adequately sampled by 2–4 mL plasma. For all samples, we calculated the detection limit (see online supplementary results) and showed that that our ability to detect ctDNA was similar in the marker-positive and marker-negative samples (see online supplementary figures S1–2 and S3A–C). The detection limit measurement furthermore showed that the sensitivity of our ctDNA assays (0.005%) would have been sufficient to detect one copy of ctDNA in the marker-negative samples, had it been present.

Serial monitoring of CEA

CEA analysis indicated relapse in 4/6 relapsing and in 0/5 non-relapsing patients, yielding a sensitivity of 67% and a specificity
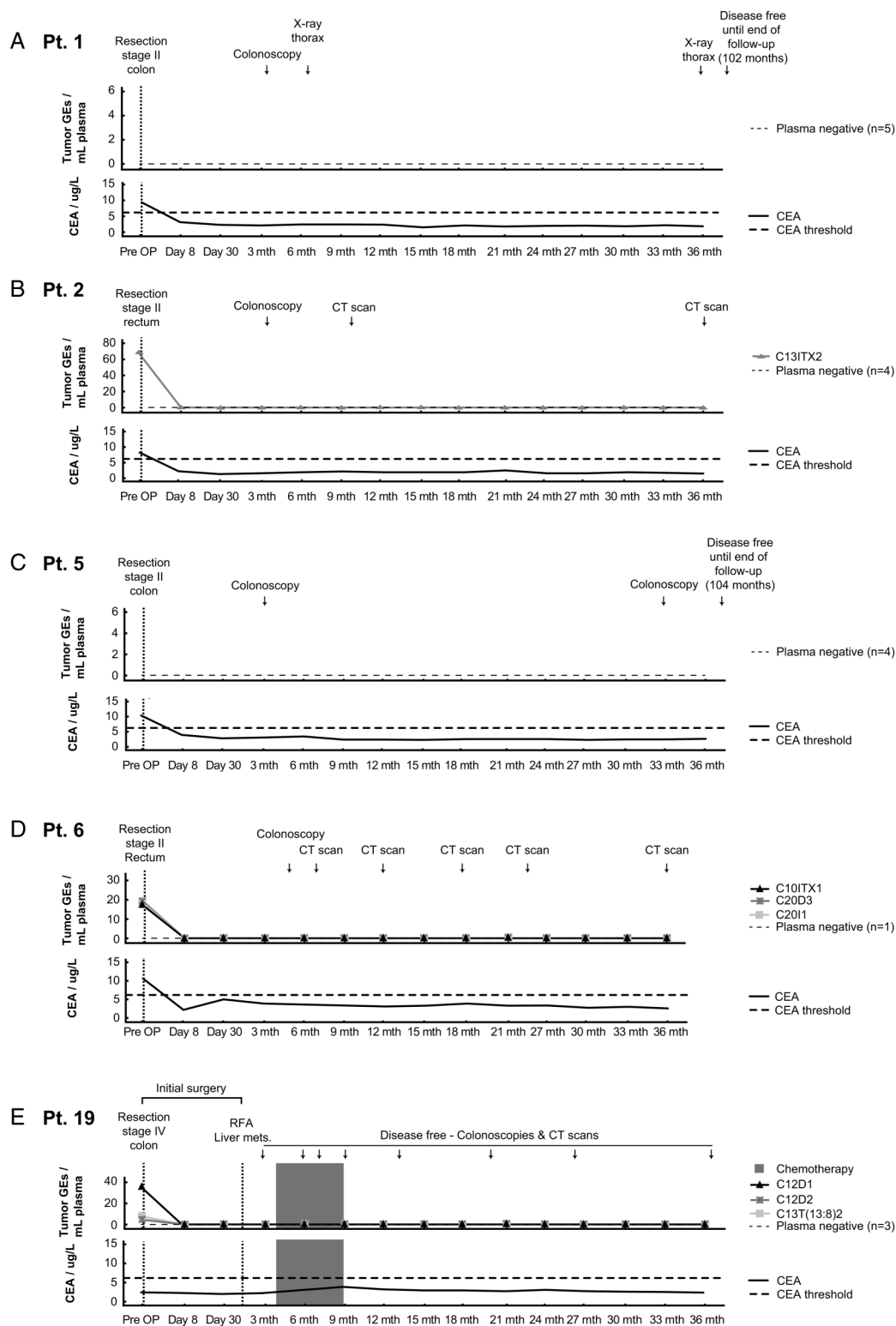
Figure 3 Clinical application of circulating tumour DNA (ctDNA) and carcinoembryonic antigen (CEA) for monitoring colorectal cancer in patients without recurrence of disease. Cell-free DNA (cfDNA) and CEA were isolated from serial plasma and serum samples, respectively. Samples were collected every third month starting prior to surgery and ending at month 36 postsurgery. TaqMan assays, amplifying patient-specific somatic structural variations identified by mate-pair sequencing, were designed and applied to cfDNA to enable quantification of the level of ctDNA. Shown are the results of monitoring the level of ctDNA (upper figure) and CEA (lower figure) in patients surgically treated with curative intend for (A) a stage II colon cancer (Pt. 1); (B) a stage II rectum cancer (Pt. 2); (C) a stage II colon cancer (Pt. 5); (D) a stage II rectum cancer (Pt. 6); (E) a stage IV colon cancer with focal metastasis in the liver. The initial treatment included a colon resection to remove the primary tumour and radiofrequency ablation (RFA) of liver metastasis (Pt. 19). The quantified levels of ctDNA are plotted as tumour genome equivalents (GEs). Vertical dotted lines indicate surgery or RFA. Grey shaded regions indicate chemotherapy. Arrows indicate radiological and molecular examinations, and they were negative unless specified otherwise. CEA threshold levels are indicated by horizontal dotted lines. Threshold values are 4 and 6 μg/L for non-smokers and smokers, respectively.

Table 3 ctDNA and CEA lead times (months) in patients with relapsing disease

Patient	Lead time* ctDNA	CEA	Δlead time ctDNA—CEA
4	15	3	12
8	7	0	7
10	11	0	11
15	12	0	12
16	13	7	6
18	2	11	−9
Average†	10	3.5	6.5

*Lead time compared with the conventional follow-up.
†ctDNA lead times are statistically different from CEA lead times (Mann–Whitney, p=0.037).
CEA, carcinoembryonic antigen; ctDNA, circulating tumour DNA.

of 100%. For the four patients where CEA indicated relapse, the average lead time was 3.5 months compared with conventional follow-up (figures 2 and 3). Moreover, the level of CEA did not reflect changes in tumour burden inflicted by surgical and therapeutic interventions (figure 2). Direct comparison of the ctDNA and CEA analyses showed that ctDNA was superior at detecting relapse and reflecting changes in tumour burden (figures 2 and 3 and table 3).

Effect of intratumour heterogeneity on SSV selection and ctDNA quantification

Studies of four tumour/metastasis pairs from patients 18, 24, 28 and 29 revealed that 35%, 91%, 63% and 23% (mean 53%) of SSVs found in the primary tumour were preserved in the metastasis, respectively (see online supplementary table S4). Therefore, it is imperative to select multiple SSVs and to select those most likely to be present in later recurrences. For the six patients with relapsing disease, 71% (17/24) of the SSV assays applied to plasma detected ctDNA. Based on this, we conclude that our SSV selection algorithm in general mitigates the potential problem of intratumour heterogeneity. In patient 18, however, only three out of six selected SSVs were detected in plasma (figure 2E, table 2), and we, therefore, made an in-depth study of the heterogeneity of the tumour and the two metastases as an illustration of the heterogeneity problem (figure 4 and see online supplementary results). Multiple samplings from primary tumour and metastases revealed spatial heterogeneity in the frequency of SSVs in the primary tumour. The metastases were spatially homogenous but each contained only a subset of SSVs found in the primary tumour. Importantly, these were SSVs found in plasma. Hence, the heterogeneity analysis unambiguously explains the plasma ctDNA findings and demonstrates the necessity for selecting multiple SSVs.

DISCUSSION

This study applied SSVs to quantify ctDNA postsurgery in curatively resected patients. Sensitivity and specificity were both 100% in terms of detecting relapse, and the average lead time was 10 months. Additionally, the quantified level of ctDNA in serial postsurgery plasma samples showed an intimate relationship with the clinical disease course. Dissimilar postintervention changes in SSVs indicate the existence of multiple lesions or cellular subclones and inform about non-radical interventions. The SSV selection algorithm presented compensated for different cellular composition of tumours and metastases, and leads us to

recommend at least three assays per patient. Although the sensitivity of ctDNA in terms of detecting relapses in this study was 100%, it is not unlikely that sensitivity will reach a level below 100% when applied to large numbers of patients. With three SSV assays per patient and 71% of the selected SSVs being present in the relapsing tumour, the probability of having three negative assays is $(1-0.71)^3$ ~2.5%. Applying five assays per patient lowers the risk of getting only negative assays to 0.2% or just two out of a thousand patients. We find that this is satisfactory and sufficient also for prospective use.

It is important to note that this was a retrospective study, and that samples were not collected with the aim to compare blood analyses to CT scans. Hence, while we in all cases detect ctDNA before the relapse was diagnosed, typically by a CT scan, the CT scans were generally not performed at the time points where blood was drawn. This means that while we in 50% of cases get an indication that ctDNA analysis is more sensitive than CT scanning, because ctDNA is detected prior to, or at the same time as, an apparently normal CT scan, we cannot assess this for the remaining 50% of patients.

The ddPCR assays designed in this study were linear across three orders of magnitude, entirely specific and capable of detecting tumour DNA even when it constituted only a very small fraction of the DNA sample, down to 0.005%. Still, we observed clinical settings where the fraction of tumour DNA in patient plasma was exceedingly small and undetectable. A recent study reported the average number of ctDNA copies per 5 mL plasma to be just 8 for stage I tumours.[15] Hence, given that we only analysed 2–4 mL of plasma, there is a risk that we misleadingly called samples negative in situations where the level of ctDNA was just very low. We confirmed that PCR inhibition was not responsible for the false-negative ctDNA samples by performing a spike-in of 10 copies of tumour DNA to ctDNA-negative plasma DNA eluate. One method to increase the amount of input material per assay is multiplexing, but the complexity of the assay design and optimisation in our hands caused the workload to increase to a level where it did not seem feasible for a clinical setting.

Some studies have reported cfDNA as an accurate measurement of tumour burden.[17 23] Generally, the studies reporting a high correlation between cfDNA and disease burden analysed plasma from metastatic patients where the level of cfDNA (and ctDNA) is higher than normal.[15 17] We hypothesise that for patients with low tumour burden, the release of ctDNA is so low that it is negligible compared with the background level of cfDNA. Indeed, this is also what we observe. This implies that ctDNA compared with cfDNA is a more sensitive marker for residual disease detection and for monitoring response to adjuvant treatment. Furthermore, our results indicate that while ctDNA is strictly tumour specific, the sources and causes to changes in the cfDNA level are numerous. The fluctuation in the cfDNA level likely reflects that it is affected by DNA released from many other sources than just the tumour cells, for example, DNA released from (A) normal cells dying due to the resection trauma, yielding a cfDNA peak at day 8 (see online supplementary figures S4, S5 and S7); (B) cells dying due to complications associated with surgery; (C) cells dying due to concomitant disease; (D) cells dying due to adjuvant chemotherapy (see online supplementary figure S7); and (E) leucocytes rupturing and releasing DNA during the handling of blood samples.

An elevated level of CEA has been reported to be a marker of residual disease with a sensitivity and specificity of approximately 80% and 70%, respectively.[24] The average lead time has been reported to be close to 5 months, but following surgery
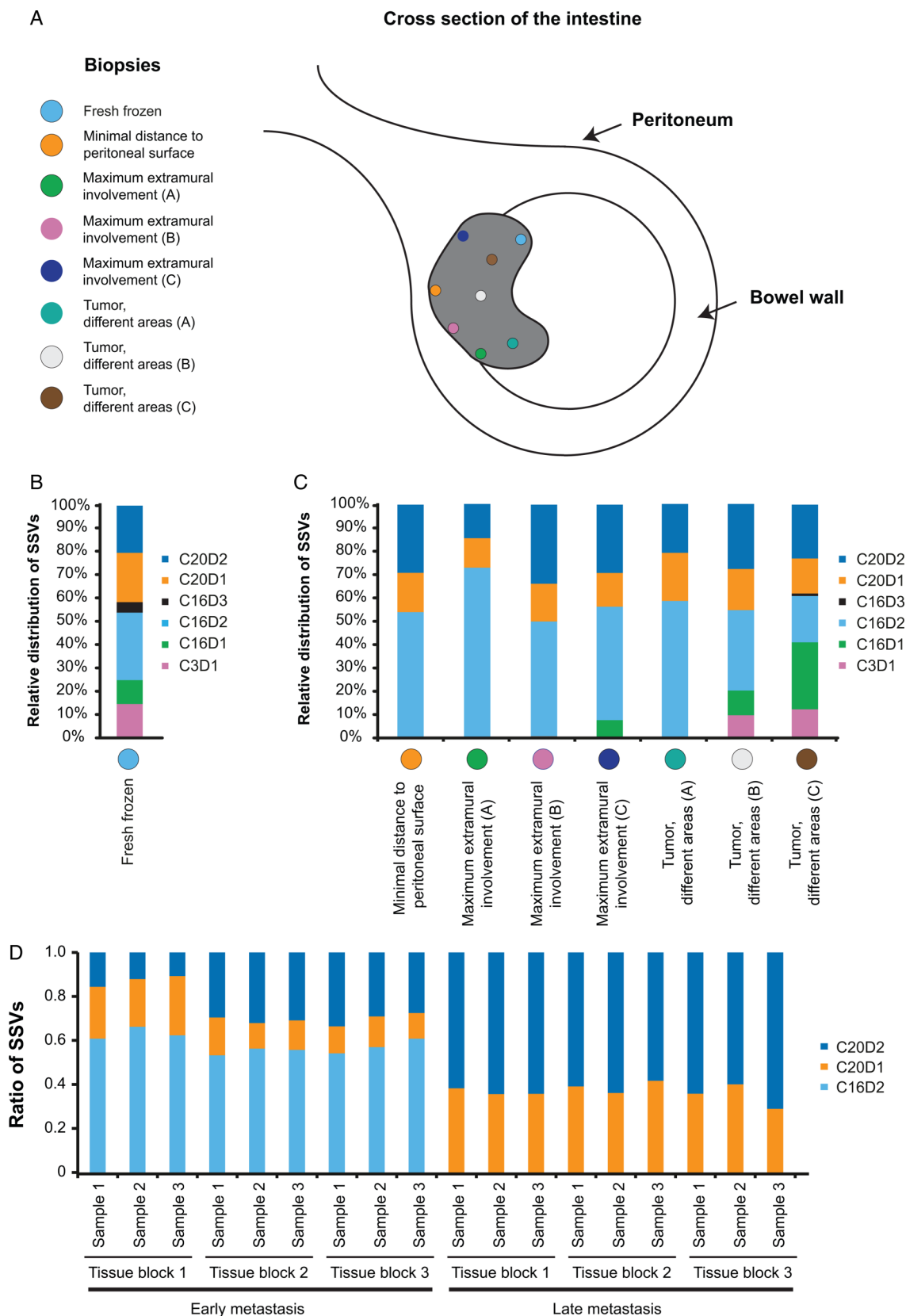
Figure 4 Heterogeneity of tumour and metastases. To investigate heterogeneity, biopsies from distinct topological locations in the primary tumour and the metastases of patient 18 were analysed with regards to somatic structural variations (SSVs). For each biopsy, the relative distribution of SSVs was calculated as the abundance of each individual SSV relative to the sum of all SSVs. (A) Schematic representation of the approximate locations of the different biopsies in the primary tumour. (B) Relative distribution of SSVs in the fresh frozen biopsy used for mate-pair sequencing and in which SSVs were originally identified. (C) Relative distributions in seven formalin fixed paraffin embedded (FFPE) biopsies sampled from different locations in the primary tumour. (D) Relative distributions of SSVs in FFPE punch biopsies taken from nine different locations in each of the early (3 months postoperative) and late (38 months postoperative) metastases. Note that only two of the examined SSVs are present in the late metastasis.

the CEA level may remain high for several weeks despite complete resection of the cancer.[24] The present results indicate that ctDNA is superior to CEA in all clinical aspects when it comes to monitoring patients with CRC.

Recent papers have shown that assays specific for point mutations in hotspot mutated genes like KRAS, BRAF and PIK3CA can be used to quantify the level of ctDNA fragments in plasma and serum in ~50% of patients with CRC.[15 25] Structural genomic alterations by contrast are present in nearly all tumours, making the approach described in the present paper generally applicable.[14] Recently, reports have shown that in the metastatic setting, where the tumour DNA constitutes a significant fraction of the total cfDNA, it is possible to identify tumour-specific alterations by both whole-genome and targeted approaches.[16 26] We foresee that future studies of ctDNA will combine the different approaches, for example, use hotspot mutations in KRAS, BRAF or PIK3CA in the patients where these are available, and then direct sequencing or an approach like the one outlined here for ~50% of patients without hotspot mutations.

The present study adds evidence indicating many promising clinical perspectives of monitoring the level of ctDNA during follow-up after curatively intended CRC surgery. The perspectives include (1) assessment of radicality of primary and secondary resections; (2) early detection of disease recurrence compared with conventional methods, thereby creating a critical window of opportunity for intervention at an early time point, where curative modalities is still an option; and (3) monitoring of response/resistance to radiation therapy, adjuvant and palliative chemotherapy in order to provide the oncologists with a guiding tool for when to stop/pause therapy, change regimen or initiate programmes aimed at identifying the location of the recurrent lesion with the possibility to perform curative surgery.

REFERENCES

1   Ferlay J, Soerjomataram I, Ervik M, et al. *Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11.* Lyon, France: International Agency for Research on Cancer, 2013. http://globocaniarcfr (accessed 14 May 2014) GLOBOCAN 2012 v1.0.
2   Rao AR, Kagan AR, Chan PM, et al. Patterns of recurrence following curative resection alone for adenocarcinoma of the rectum and sigmoid colon. *Cancer* 1981;48:1492–5.
3   Bohm B, Schwenk W, Hucke HP, et al. Does methodic long-term follow-up affect survival after curative resection of colorectal carcinoma? *Dis Colon Rectum* 1993;36:280–6.
4   Jeffery M, Hickey BE, Hider PN. Follow-up strategies for patients treated for non-metastatic colorectal cancer. *Cochrane Database Syst Rev* 2007;(1):CD002200.
5   Renehan AG, Egger M, Saunders MP, et al. Impact on survival of intensive follow up after curative resection for colorectal cancer: systematic review and meta-analysis of randomised trials. *BMJ* 2002;324:813.
6   National Comprehensive Cancer Network. Clinical Practice Guidelines in Oncology: Colon Cancer Version 2. 2014. 11/2013.
7   Labianca R, Nordlinger B, Beretta GD, et al. Early colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013;24 (Suppl 6):vi64–72.
8   Meyerhardt JA, Mangu PB, Flynn PJ, et al. Follow-up care, surveillance protocol, and secondary prevention measures for survivors of colorectal cancer: American Society of Clinical Oncology clinical practice guideline endorsement. *J Clin Oncol* 2013;31:4465–70.
9   Diehl F, Li M, Dressman D, et al. Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc Natl Acad Sci U S A* 2005;102:16368–73.
10 Diehl F, Schmidt K, Choti MA, et al. Circulating mutant DNA to assess tumor dynamics. *Nat Med* 2008;14:985–90.
11 Nawroz H, Koch W, Anker P, et al. Microsatellite alterations in serum DNA of head and neck cancer patients. *Nat Med* 1996;2:1035–7.
12 The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012;487:330–7.
13 Seshagiri S, Stawiski EW, Durinck S, et al. Recurrent R-spondin fusions in colon cancer. *Nature* 2012;488:660–4.
14 Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. *Science* 2013;339:1546–58.
15 Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med* 2014;6:224ra24.
16 Dawson SJ, Tsui DW, Murtaza M, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N Engl J Med* 2013;368:1199–209.
17 Spindler KL, Pallisgaard N, Vogelius I, et al. Quantitative cell-free DNA, KRAS, and BRAF mutations in plasma from patients with metastatic colorectal cancer during treatment with cetuximab and irinotecan. *Clin Cancer Res* 2012;18:1177–85.
18 Korn JM, Kuruvilla FG, McCarroll SA, et al. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet* 2008;40:1253–60.
19 Chen K, Wallis JW, McLellan MD, et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods* 2009;6:677–81.
20 Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* 2013;14:178–92.
21 Pallisgaard N, Andersen RF, Spindler KLG, et al. MC13-0088 Controls for quantification of total cell free DNA qPCR analyses in plasma. *Eur J Cancer* 2013;49(Suppl 4):S36.
22 Andersen CL, Lamy P, Thorsen K, et al. Frequent genomic loss at chr16p13.2 is associated with poor prognosis in colorectal cancer. *Int J Cancer* 2011;129:1848–58.
23 Frattini M, Gallino G, Signoroni S, et al. Quantitative and qualitative characterization of plasma DNA identifies primary and recurrent colorectal cancer. *Cancer Lett* 2008;263:170–81.
24 Duffy MJ. Carcinoembryonic antigen as a marker for colorectal cancer: is it clinically useful? *Clin Chem* 2001;47:624–30.
25 Thierry AR, Mouliere F, El Messaoudi S, et al. Clinical validation of the detection of KRAS and BRAF mutations from circulating tumor DNA. *Nat Med* 2014;20:430–5.
26 Leary RJ, Kinde I, Diehl F, et al. Development of personalized tumor biomarkers using massively parallel sequencing. *Sci Transl Med* 2010;2:20ra14.