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]

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2014-308859).

[1]Department of Molecular Medicine, Aarhus University Hospital, Aarhus, Denmark
[2]Department of Surgical Gastroenterology, Aarhus University Hospital, Aarhus, Denmark
[3]Institute of Pathology, Aarhus University Hospital, Aarhus, Denmark
[4]Department of Surgical Gastroenterology, University of Copenhagen, Hvidovre Hospital, Hvidovre, Denmark
[5]Department of Surgery P, Aarhus University Hospital, Aarhus, Denmark
[6]Department of Clinical Biochemistry, Vejle Hospital, Vejle, Denmark

Correspondence to
Claus L Andersen, Department of Molecular Medicine, Aarhus University Hospital, Aarhus, Denmark; cla@clin.au.dk

TR and LVS contributed equally.

Received 19 November 2014
Revised 22 December 2014
Accepted 7 January 2015
Published Online First 5 February 2015

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

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

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	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.

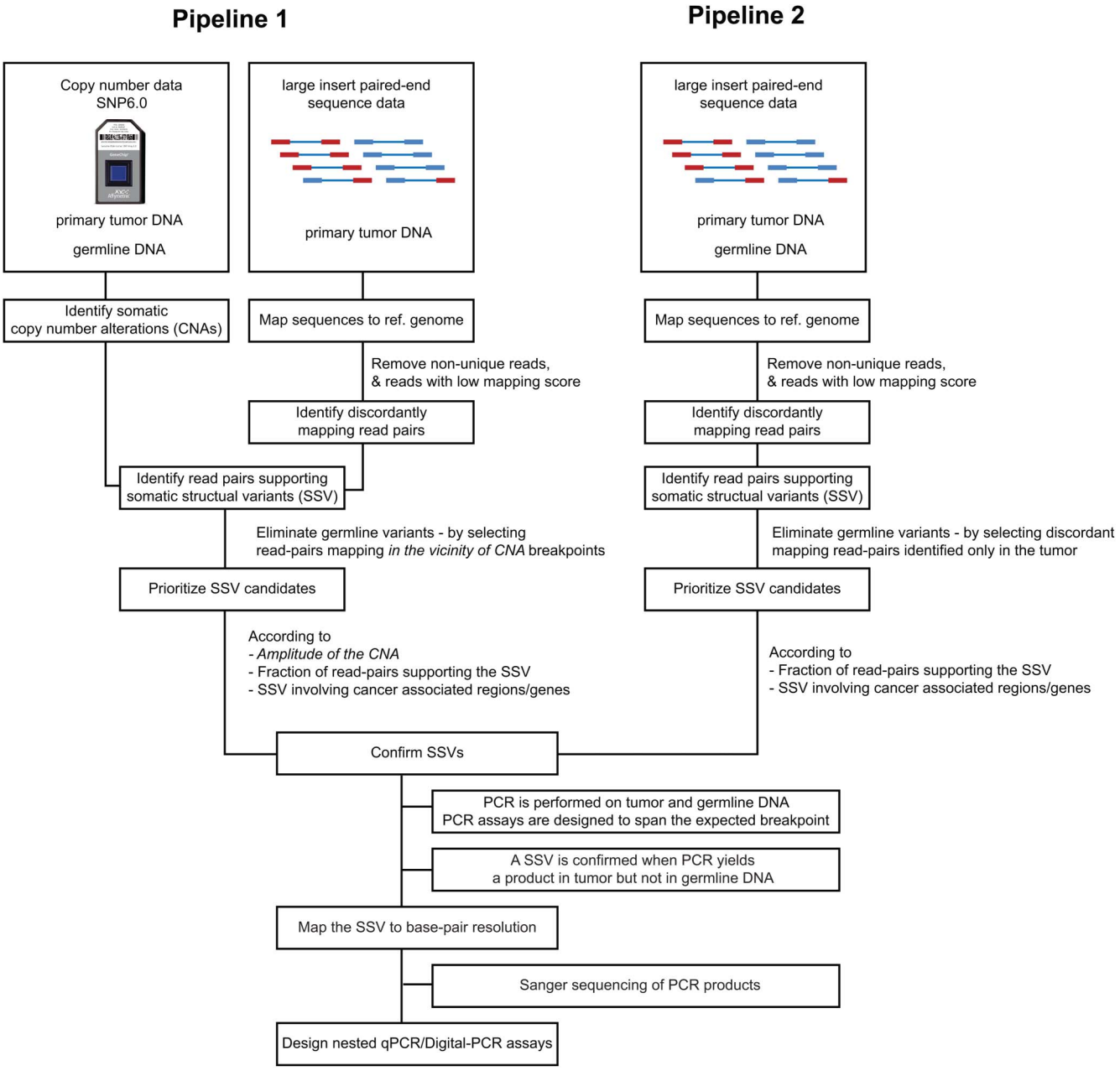


Figure 1 Pipelines for identification of somatic structural variations (SSVs). Two different pipelines for identification of SSVs in the primary tumour were applied. Both pipelines were performed without knowledge of patient outcome. In pipeline 1, only the tumour was mate-pair sequenced and used for 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,

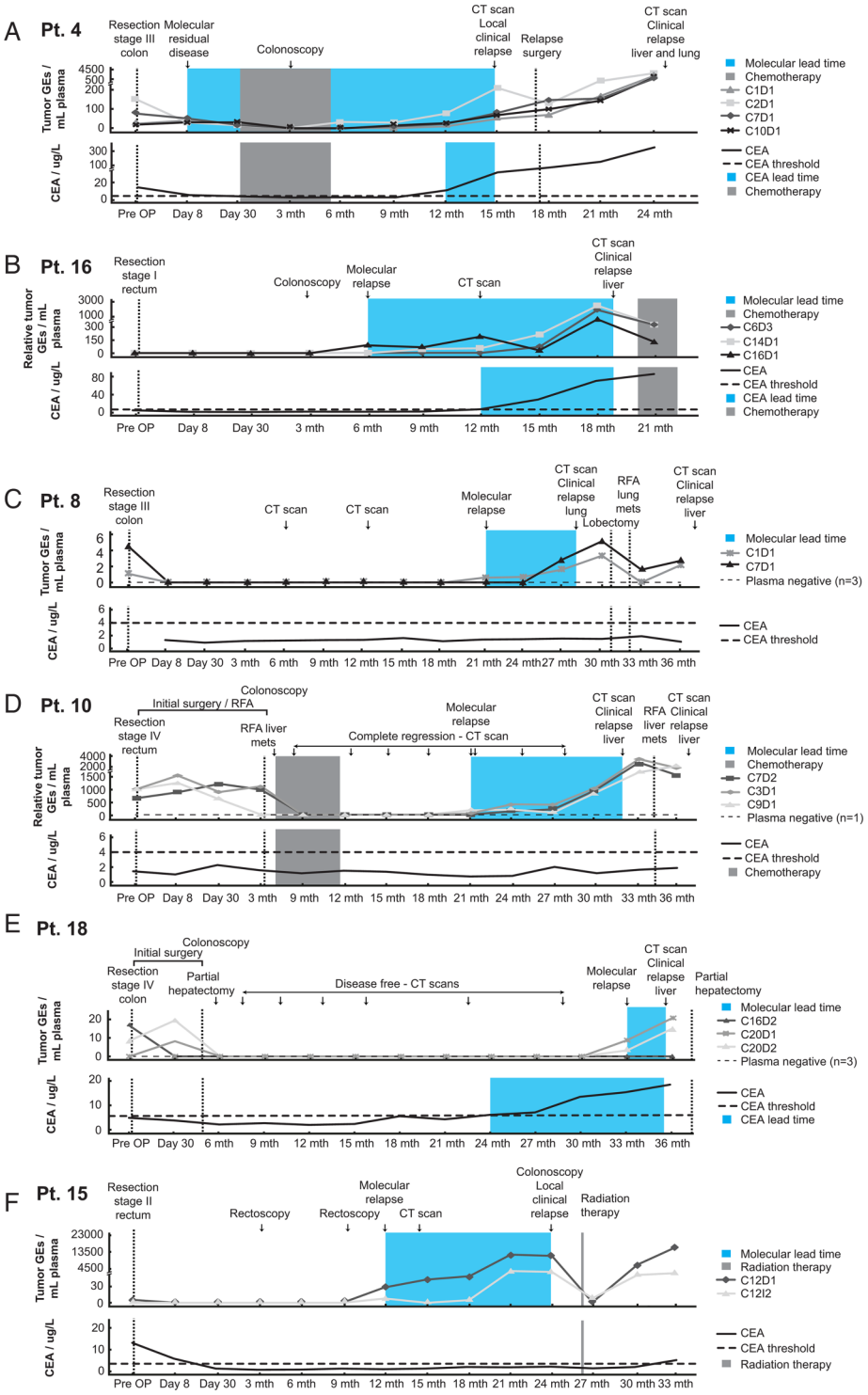


Figure 2 Clinical application of circulating tumour DNA (ctDNA) and carcinoembryonic antigen (CEA) for monitoring colorectal cancer in patients with residual disease and/or recurrence. Cell-free DNA (cfDNA) and CEA were isolated from serial plasma and serum samples, respectively. Samples were collected 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 intrachromosomal 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

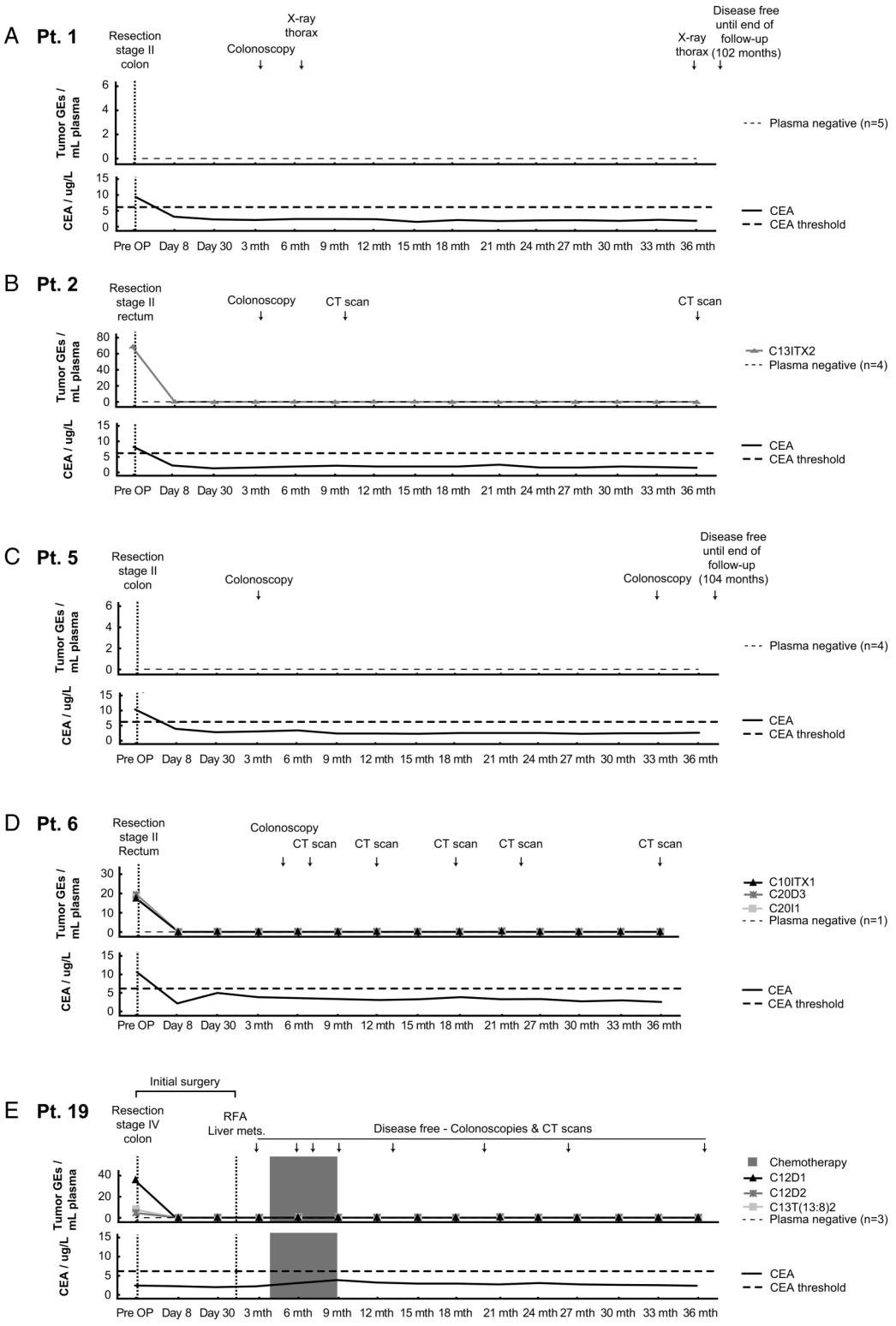


Figure 3 Clinical application of circulating tumour DNA (ctDNA) and carcinoembryonic antigen (CEA) for monitoring colorectal cancer in patients without recurrence of disease. Cell-free DNA (cfDNA) and CEA were isolated from serial plasma and serum samples, respectively. Samples were collected every third 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

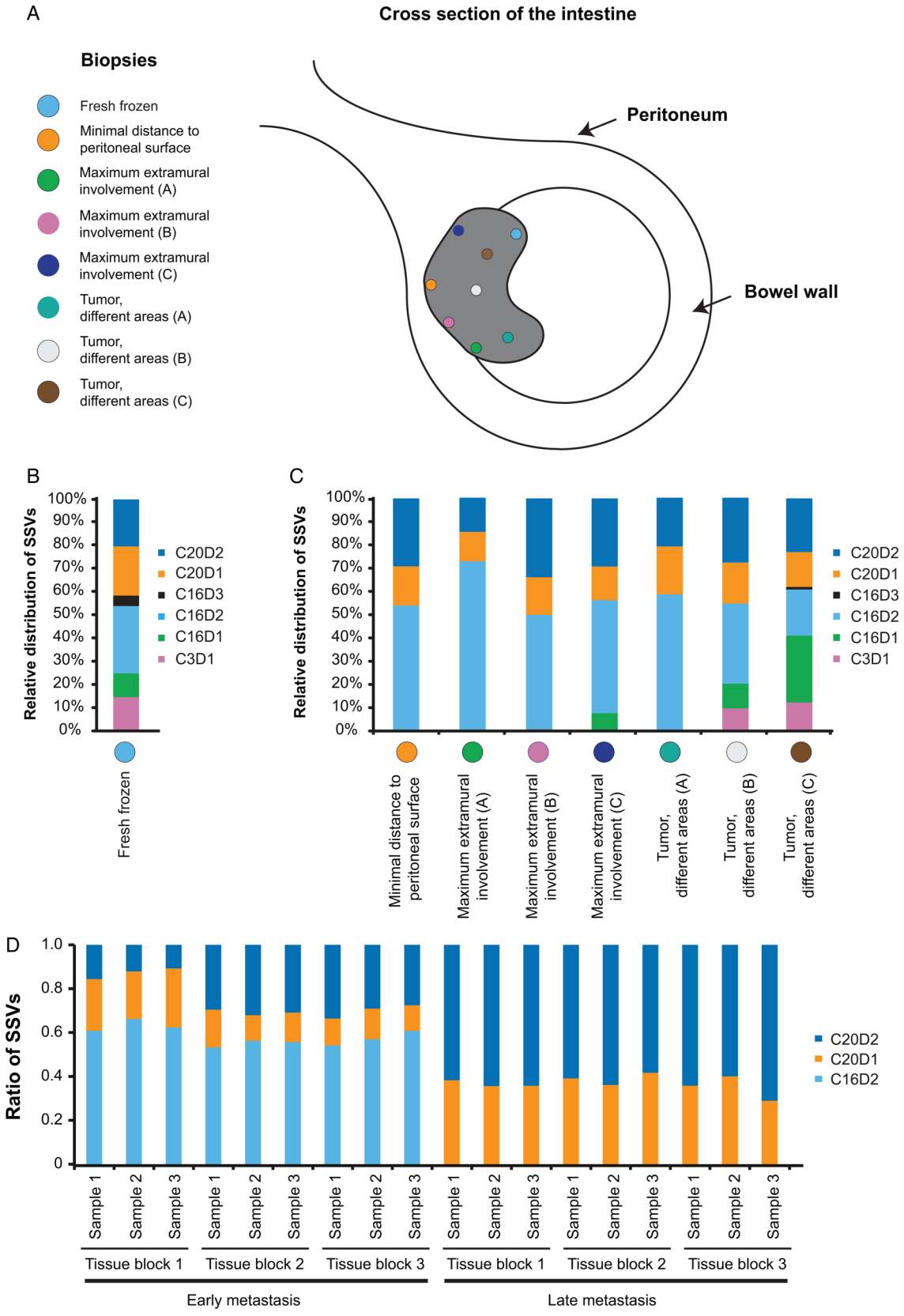


Figure 4 Heterogeneity of tumour and metastases. To investigate heterogeneity, biopsies from distinct topological locations in the primary tumour and the metastases of patient 18 were analysed with regards to somatic structural variations (SSVs). For each biopsy, the relative distribution of SSVs was 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.

Acknowledgements The authors thank the study participants and staff at the hospitals that collected samples for this study (Aarhus University Hospital and Hvidovre Hospital) and Bente Devantié, Birgitte Trolle, Hanne Steen, Karen Bihl, Margaret Gellet, Pamela Celis and Louise Nielsen for excellent technical assistance. The Danish CancerBiobank (DCB) is acknowledged for biological material and for the data information regarding handling and storage.

Contributors Study concept and design: CLA. Acquisition of data: TR, LVS, RT, HT, IN, A-SK, FVM, KS, SH-D, HJN, SL and NP. Analysis and interpretation of data: TR, LVS, SV, IN, PL, KS, JSP and CLA. Drafting of the manuscript: TR and CLA. Critical revision of the manuscript: TR, LVS, SV, IN, HJN, TFØ and CLA.

Funding This work was supported by The John and Birthe Meyer Foundation, The Danish Council for Independent Research, the Lundbeck Foundation, the Danish Cancer Society, The Toyota Foundation, the Novo Nordisk Foundation, the University of Aarhus and The Danish Ministry of the Interior and Health.

Competing interests None.

Patient consent Obtained.

Ethics approval Research protocols were approved by the Central Denmark Region Committees on Biomedical Research Ethics.

Provenance and peer review Not commissioned; externally peer reviewed.

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.

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. Estimates of the total number of cfDNA template copies per ddPCR reaction (dark gray bars) and DNA purification efficiency (dotted line) (CPP1 control) for each plasma sample for patients with residual disease and/or recurrence. (A-F) cfDNA and purification efficiency estimates for pt. 4, 8, 16, 10 18, and 15 respectively. The cfDNA template copies were calculated as the average of the absolute template copy numbers determined using two reference ddPCR assays targeting reference regions on chromosomes 3 and 7. These cfDNA quantities are used to estimate the minimal ctDNA/cfDNA ratio detectable for each sample. For a sample with 500 cfDNA template copies per reaction a positive detection of ctDNA is unlikely if the actual ctDNA/cfDNA fraction in that sample is less than 1/500 (0.2%). We use the measure to estimate the likelihood of a given plasma sample to be truly negative if our ctDNA specific ddPCR assays do not give a signal. We also use it to evaluate if the sensitivity and specificity of our ctDNA ddPCR assays are sufficient for detection of the ctDNA present in a sample, if any. Light grey bars indicate samples that tested positive using the ddPCR lymphocyte DNA contamination assay. In such samples our QC approach over-estimates the minimal ctDNA/cfDNA ratio detectable. ctDNA positive samples are indicated by an asterisk.

Supplementary Figure 2. Estimates of the total number of cfDNA template copies per ddPCR reaction (dark gray bars) and DNA purification efficiency (CPP1 control) (dotted line) for each plasma sample for patients with no residual disease and/or recurrence. (A-E) cfDNA and purification efficiency estimates for pt. 1, 2, 5, 6, and 19 respectively. The cfDNA template copies were calculated as the average of the absolute template copy numbers determined using two reference ddPCR assays targeting reference regions on chromosomes 3 and 7. These cfDNA quantities are used to estimate the minimal ctDNA/cfDNA ratio detectable for each sample. For a sample with 500 cfDNA template copies per reaction a positive detection of ctDNA is unlikely if the actual

ctDNA/cfDNA fraction in that sample is less than 1/500 (0.2%). We use the measure to estimate the likelihood of a given plasma sample to be truly negative if our ctDNA specific ddPCR assays do not give a signal. We also use it to evaluate if the sensitivity and specificity of our ctDNA ddPCR assays are sufficient for detection of the ctDNA present in a sample, if any. ctDNA positive samples are indicated by an asterisk.

Supplementary Figure 3. Boxplot showing genome equivalents of cell free DNA analyzed per SSV. (A) Boxplot of individual patients without residual disease and/or recurrence. (B) Boxplot of individual patients with residual disease and/or recurrence. (C) Boxplot of patients with or without residual disease and/or recurrence.

Supplementary Figure 4. Clinical application of cfDNA for monitoring colorectal cancer patients in patients with residual disease and/or recurrence. cfDNA was isolated from serial plasma samples collected every third month starting prior to surgery and ending at month 36 post-surgery or in relation to recurrence of disease. Shown are the results of monitoring the level of cfDNA 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 tumor and radio frequency ablation 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 tumor 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 cfDNA are plotted as cfDNA 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. Blue shaded regions indicate lead time decided by detection of ctDNA. * QC indicates contamination with lymphocyte DNA.

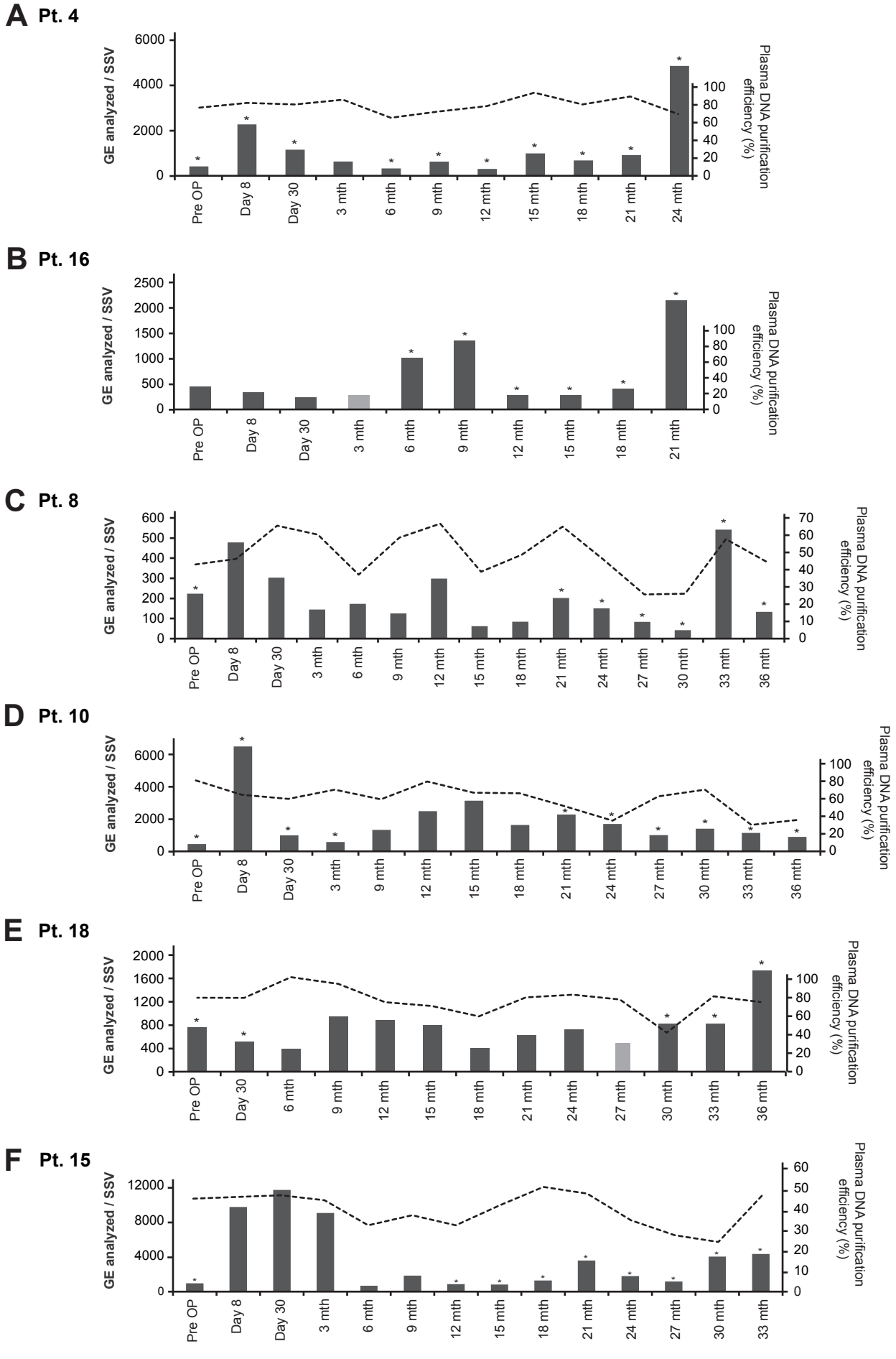
Supplementary Figure 5. Clinical application of cfDNA for monitoring colorectal cancer patients in patients without residual disease and/or recurrence. cfDNA was isolated from serial plasma samples collected every third month starting prior to surgery and ending at month 36 post-surgery or in relation to recurrence of disease. Shown are the results of monitoring the level of cfDNA and ctDNA 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 tumor and radiofrequency ablation of liver metastasis (Pt. 19). The quantified levels of cfDNA are plotted as cell-free GEs. Data are only shown for informative assays (See Supplementary Table 4 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.

Supplementary Figure 6. Correlation between cfDNA and ctDNA estimates. The correlation between cfDNA and ctDNA levels was analyzed by scatter plot and Spearman correlation for (A) All plasma samples, (B) All ctDNA positive plasma samples, (C) All ctDNA positive samples with less than 63.9 ctDNA GEs/mL plasma, (D) All ctDNA positive samples with more than 63.9 ctDNA GEs/mL plasma.

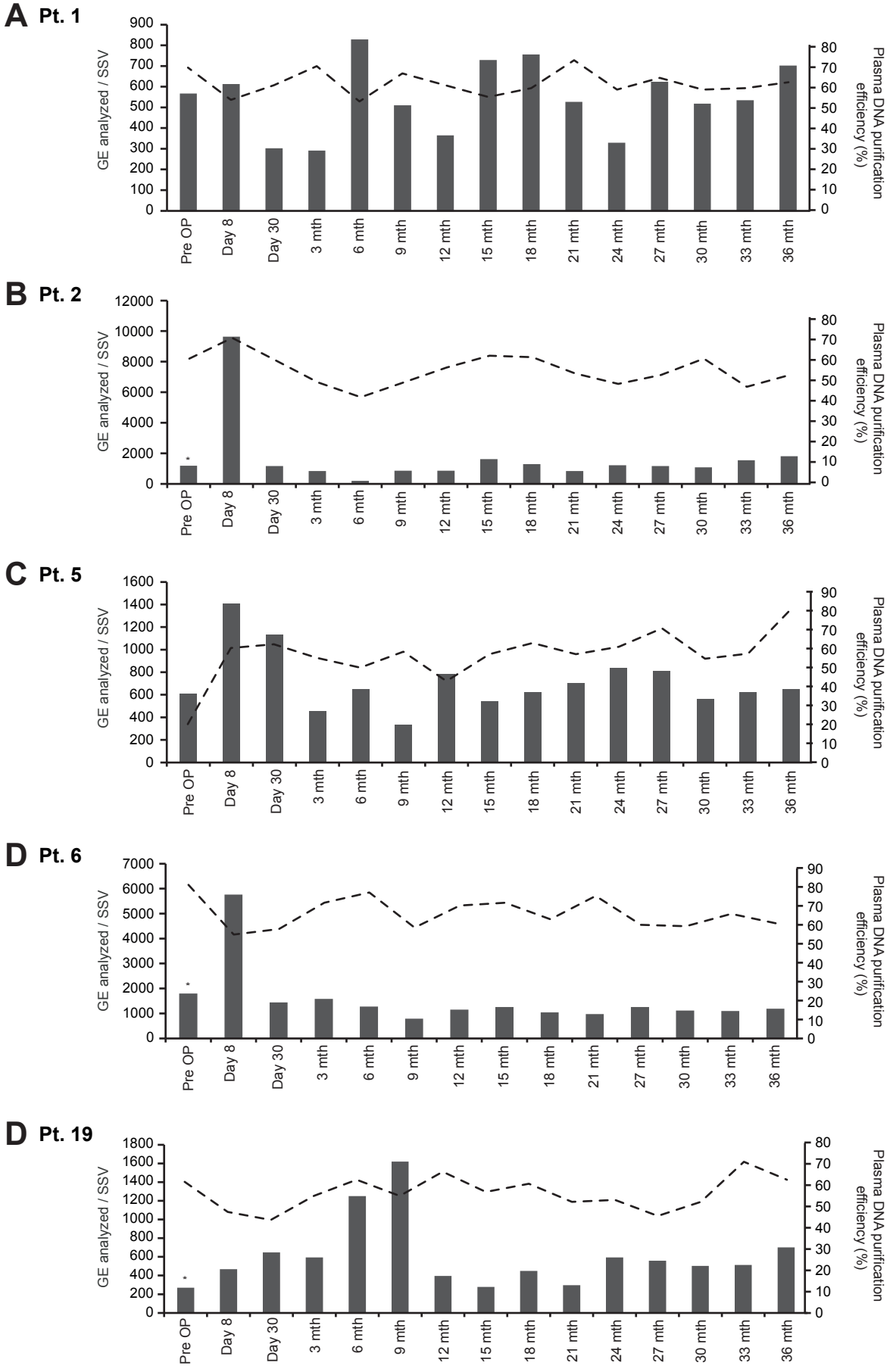
Supplementary Figure 7. cfDNA levels in ctDNA negative plasma samples. cfDNA GE per mL plasma is plotted for all ctDNA negative samples according to time of sampling. Black dots indicate plasma samples with less than 5000 GEs cfDNA. Green dots indicate samples with more than 5000 GEs cfDNA, but no obvious course. Orange dots indicate samples with more than 5000 GEs cfDNA in patients with complications related to surgery. Purple dots indicate samples with more than 5000 GEs cfDNA in patients with trauma from the resection at day 8 post OP. Blue dots

indicate samples with more than 5000 GEs cfDNA and from a patient suffering from CLL. Gray dots indicate samples with more than 5000 GEs cfDNA due to chemotherapy.

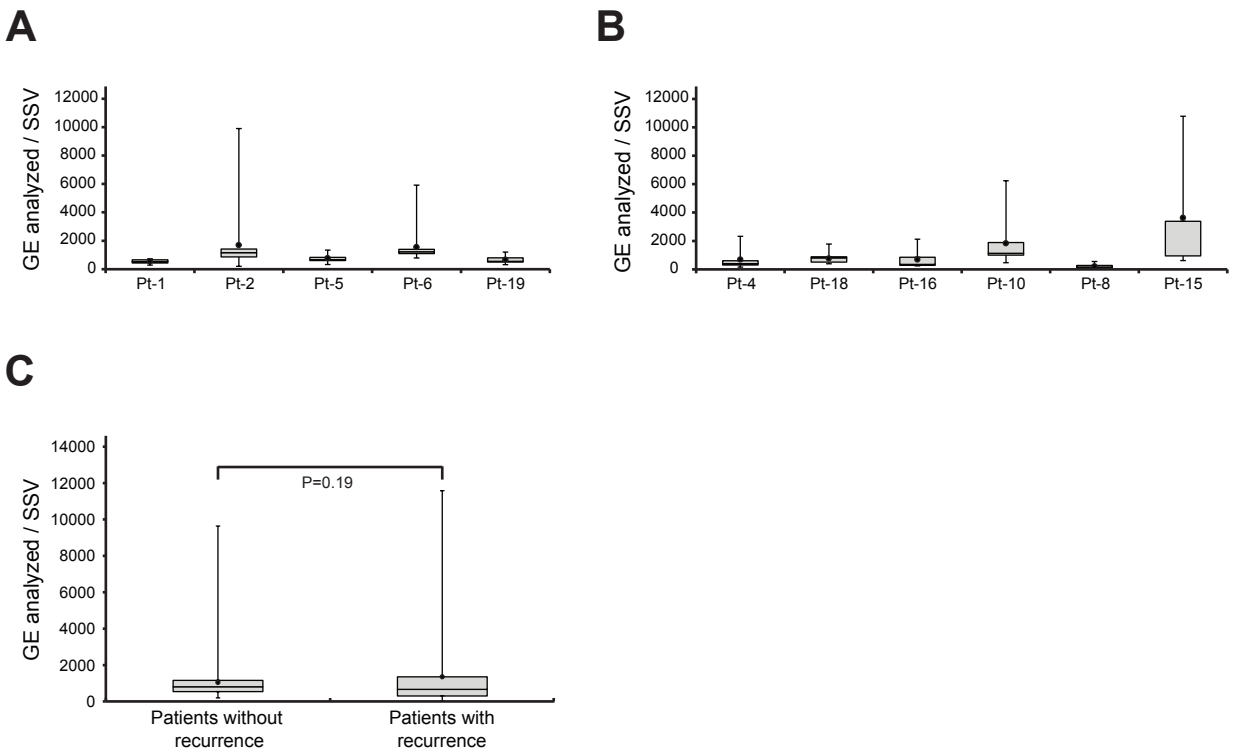
Supplementary Figure 1



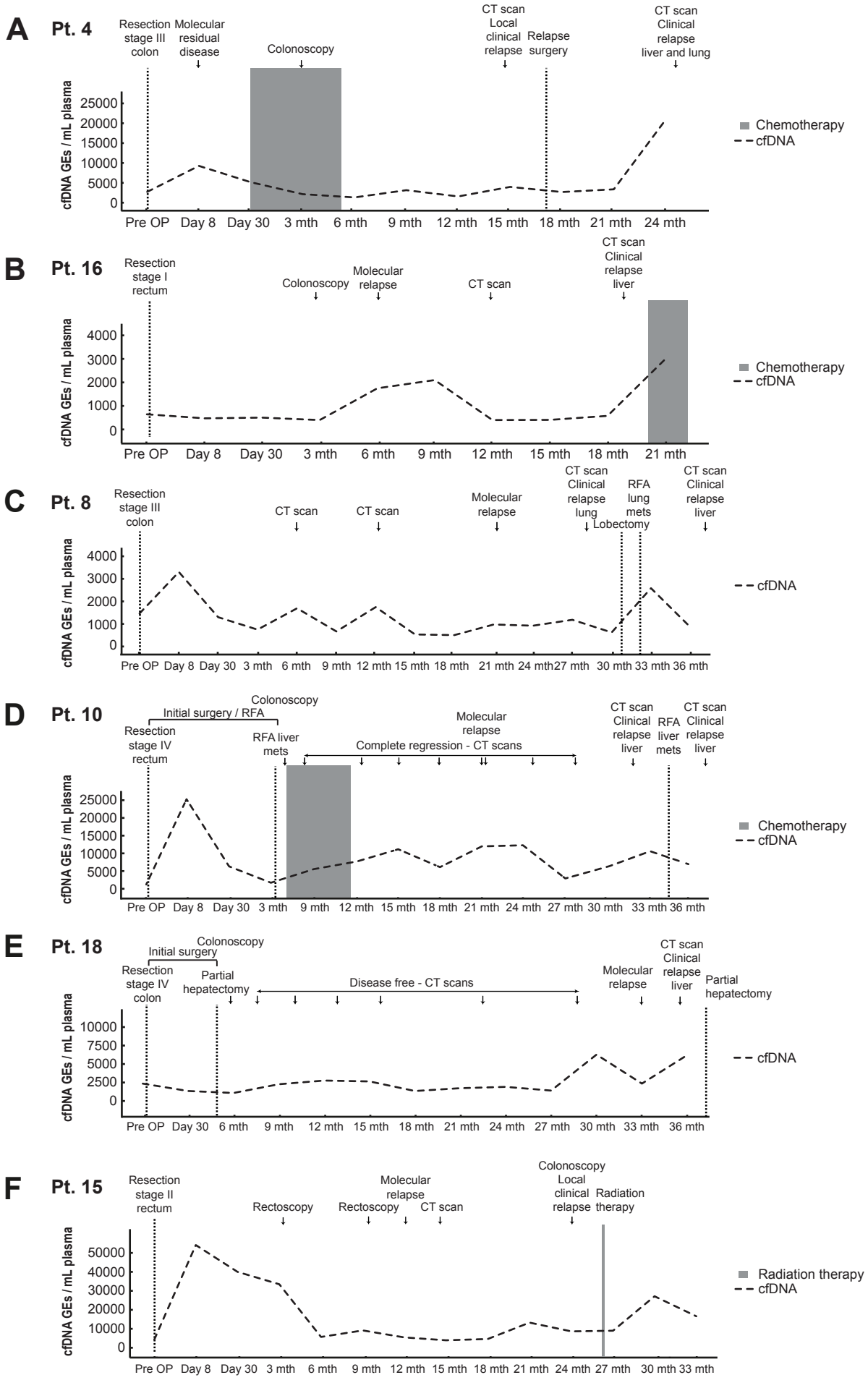
Supplementary Figure 2



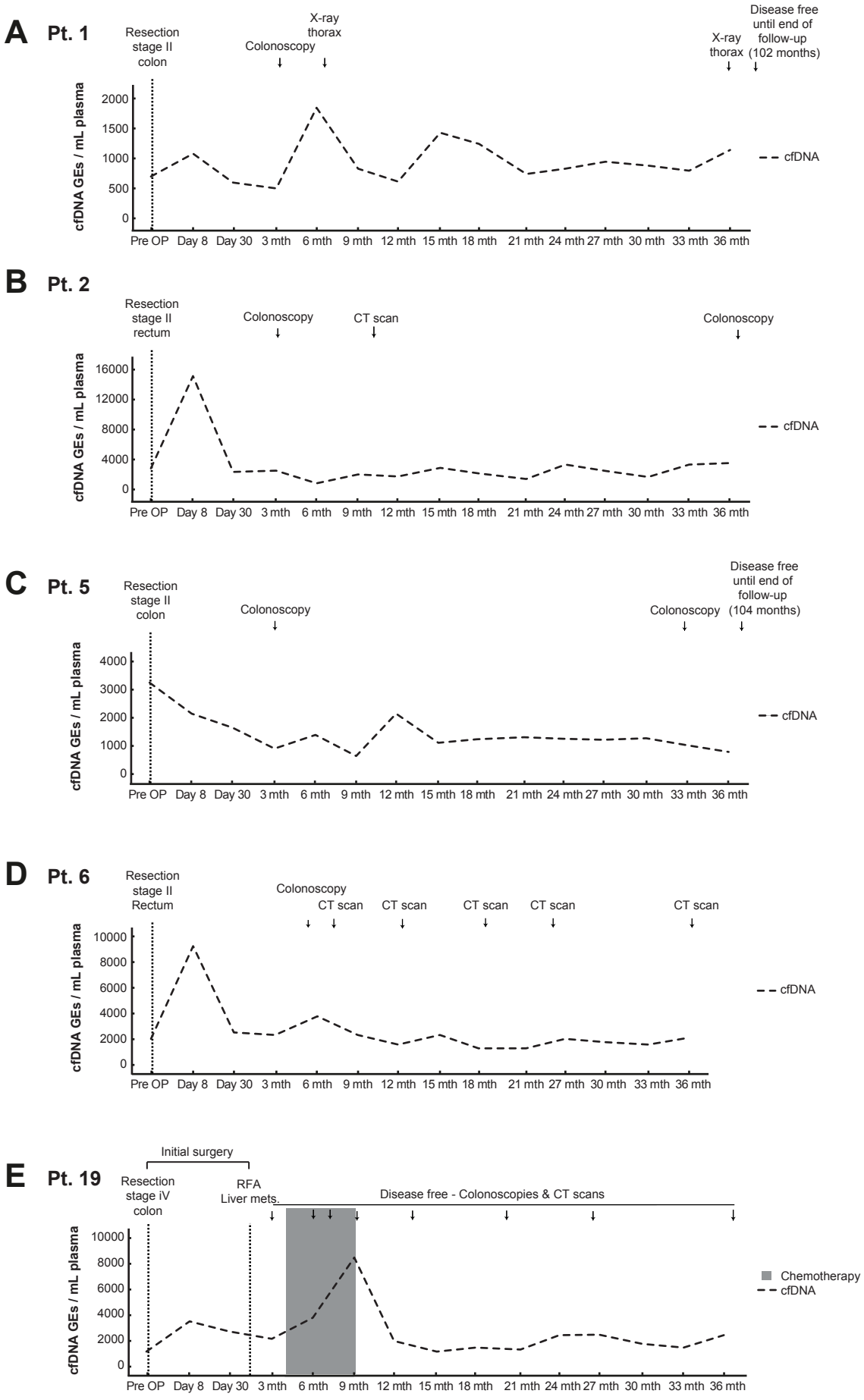
Supplementary Figure 3



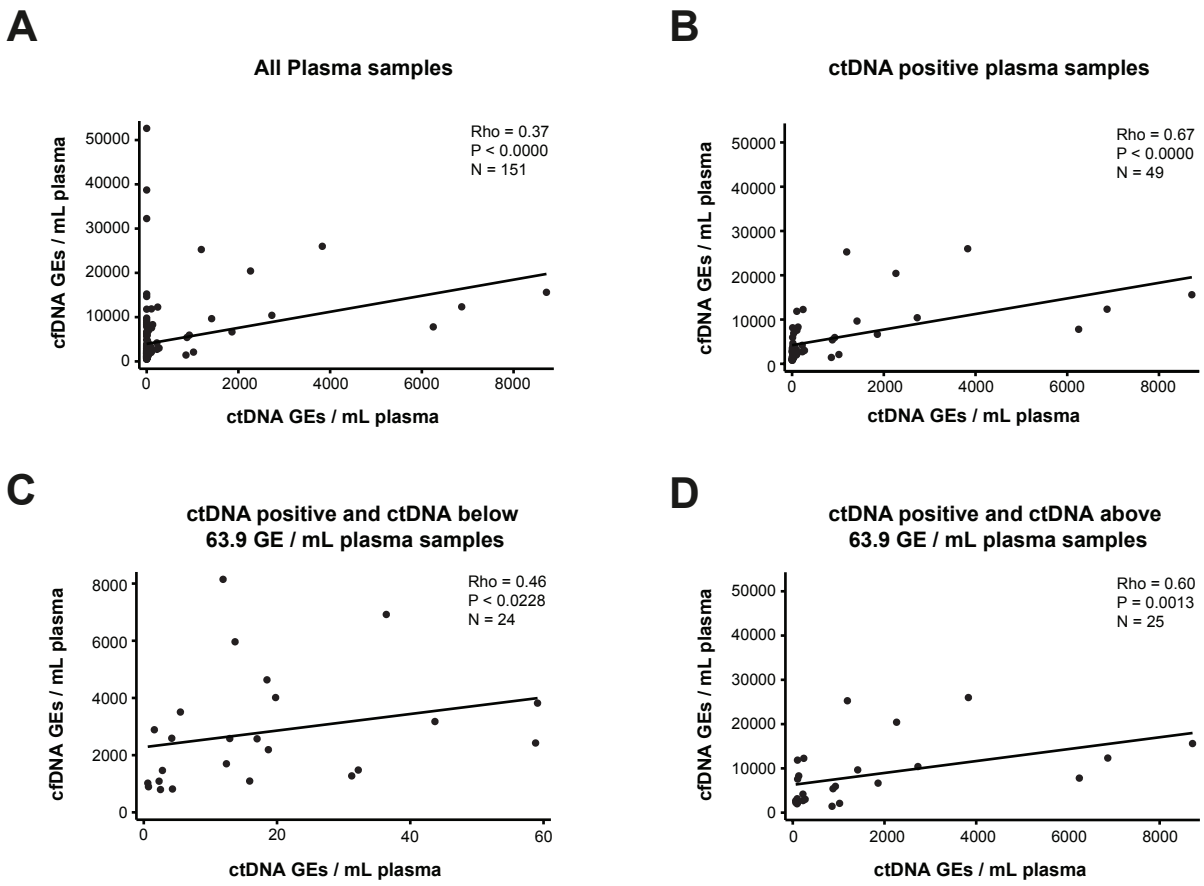
Supplementary Figure 4



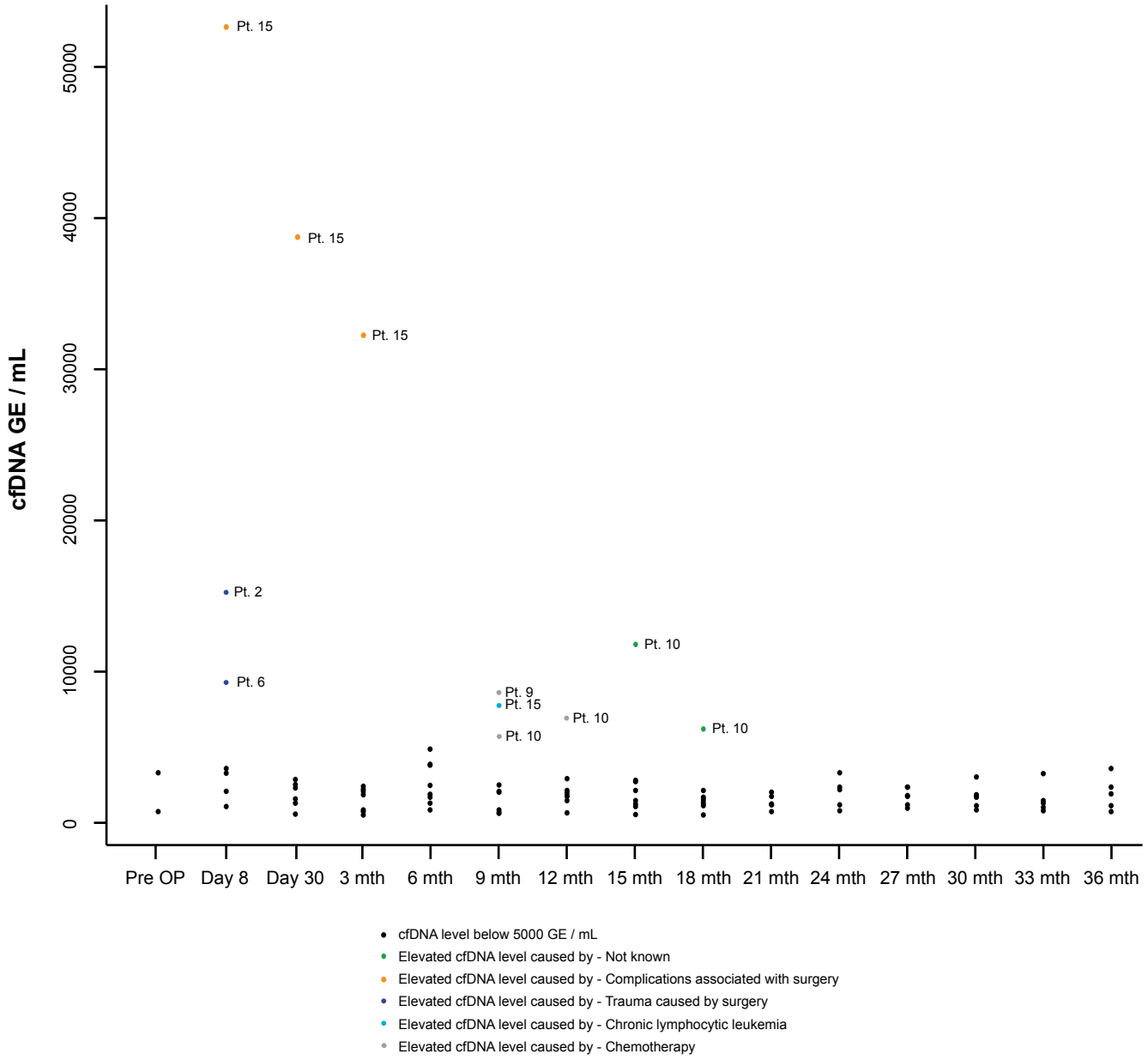
Supplementary Figure 5



Supplementary Figure 6



Supplementary Figure 7



SUPPLEMENTARY DATA

Supplementary Methods

Identification of somatic copy number alterations by SNP-array data

Matched germline and tumor DNA were profiled using SNP6.0 arrays (Affymetrix). Initial QC of the arrays was performed using the Birdsuite software.[1] The QC acceptance criteria were contrast QC >40, SNP call % >97, and MAPD <0.4. Tumor specific copy number alterations (CNAs) were derived from each tumor/normal pair by running the "CRMA (v2): Paired total copy number analysis" from the aroma.cn package in R.[2] The package is available at aroma-project.org.

Illumina mate pair libraries and next generation sequencing

Initially, we used the Illumina Mate Pair Library Preparation Kit v2 together with the TruSeq ™ DNA Sample Preparation Kit to allow indexing of mate pair libraries. Briefly, 2.5 ug of high molecular weight genomic DNA (gDNA) was fragmented by Covaris Adaptive Focused Acoustics™ (AFA) sonication device (S2, Covaris, Inc.) to a fragment size of 2000-5000 bp. (Duty cycle 20%, intensity 0.1, 1000 cycles burst, 5 minutes). Sonication was performed in AFA miniTUBE (Part # 520066). Samples were speedvaced and the fragment size analyzed on an Agilent DNA chip (Bioanalyzer). Following Biotinylation, according to the Illumina protocol, the samples were indexed by performing end repair, A-Tailing, and adapter ligation as described in the TruSeq™ protocol. Finally the libraries were enriched by 18 cycles of PCR. The DNA concentration of the libraries was evaluated by Q-PCR (KAPA Library Amplification Kit, KK2611; KAPABIOSYSTEMS) and the insert size distribution was measured by Agilent DNA 1000 Analyzer Chip. Lately, we used the Nextera Mate Pair Sample Preparation Kit (FC-132-1001, Illumina) that utilizes a gel free protocol and reduces the input to 1 ug of genomic DNA. We adjusted the concentration of each library to 2 nM and prepared clusters on Illumina paired end flow

cells using the manufacturer's instructions and the appropriate Illumina cluster and sequencing kits. Illumina sequencing was performed on the Illumina HiSeq2000 platform. Here, we employed the sequencing kits TruSeq SBS Kit v3 - HS (200-cycles) (FC-401-3001, Illumina) and TruSeq PE Cluster Kit v3 - cBot - HS (PE-401-3001, Illumina) to generate 2x101 bp paired end sequencing. Average sequence read depth/physical read depth is listed in Supplementary Table 2.

Data Analysis

Fastq files were prepared with CASAVA (v1.8.2) and quality checked using FastQC and FastqScreen (http://www.bioinformatics.babraham.ac.uk/projects/). For data generated by the Illumina Mate Pair Library Preparation Kit v2 the read length was cut down to 50 base pairs to minimize the number of chimeric reads as a result of reading through the circulation ligation point. For the Nextera DNA Sample Preparation Kit the internal transposon adapters were removed using Cutadapt (https://code.google.com/p/cutadapt/) and Illumina adapters were removed using AdapterRemoval (v1.5).[3] Next the reads were reverse complemented using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) and mapped to hg19 using BWA (v0.6.1-r104).[4] Insert size and PCR and optical duplicates were marked in each library using the Picard package v1.88 (http://picard.sourceforge.net). To clean the data, unmapped reads, PCR duplicates, and read pairs with insert sizes below 1000 bp were removed. BreakDancer[5] was applied to the final alignments with a mapping quality cut-off at 10 for the germline samples and 35 for tumor and metastasis samples to identify discordant mapping reads and to annotate them according to the underlying type of structural variant e.g. deletion (DEL), intra chromosomal translocation (ITX). Inter chromosomal translocation (CTX), inversion (INV), and insertion (INS). Bed-files were produced to visually inspect and compare the BreakDancer output with the SNP array copy number data in the IGV browser.[6] To check for sample mixup, SNPs were called using samtools mpileup and hierarchical clustering of all libraries was performed using the mitochondrial SNPs.

Identification of somatic structural variants

We applied two different pipelines to identify SSVs (Figure 1). Pipeline 1 was based on SNP array data to identify somatic copy number alterations (CNA) by comparing tumor and blood DNA. CNAs were taken further if they were supported by 15 or more probes, had a genomic size above 20,000 bp, and had a mean log2 copy number ratio amplitude >|0.2|. The somatic CNAs were compared to the tumor mate pair analysis to identify read pairs supporting the SNP findings. The number of read pairs supporting each CNA was logged. Four to eight SSVs from each patient were selected for validation based on the amplitude of the SNP analysis, the number of supporting reads in the mate pair analysis, and the possible biological function of the involved genes. SSVs with a high absolute log2 copy number ratio amplitude and many supporting reads in the mate pair analysis was preferred, as this indicate that many tumor cells contain this allele, or the region is amplified in all or some of the tumor cells. SSVs affecting known drivers of CRC tumorigenesis were also preferred as these are likely to have occurred early in tumor development and therefore to be present in the majority of tumor cells and subsequent metastasis. In pipeline 2, the SSVs were identified by analysis of matched germline DNA, but otherwise the selection criteria were the same as in pipeline 1. Potential SSVs where evaluated by PCR using tumor and germline DNA. PCR products were visualized on an agarose gel and Sanger sequenced across the SSV junction to annotate the SSV breakpoint at base-pair resolution.

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) with O/N incubation with proteinase K at 56˚C and 550 rpm, from 2-4 mL plasma by QIAamp DNA Blood Midi Kit (51185), modified to be able to allow two mL of plasma on one column, or QIAamp Circulating Nucleic Acid Kit (55114) according to the manufactures instruction.

Amplification of SSVs by multiplex Nested PCR

For the analysis of patients 10 and 16, twelve cycles of Nested PCR were carried out with seven sets of primers (Supplementary Table 3) using 90% the DNA purified from the plasma samples (the remaining 10% were used to estimate CPP1, cfDNA quantity, and leucocyte DNA contamination) as template in a final volume of 50 µl and the following final concentrations: primer mix 0.5 µM, dNTP 0.2 mM, PCR buffer 1x, $MgCL_2$ 1.5 mM, Tempase enzyme 0,1U/µl. The Nested product was diluted between four and two hundred times before being used as template in ddPCR.

SUPPLEMENTARY RESULTS

Detection limit

For all samples we stablished a procedure for assessing the minimal ctDNA/cfDNA ratio detectable in a given sample. For a sample with no leucocyte DNA contamination and for which cfDNA purification was 100% efficient the minimal detectable fraction (the detection limit) is 1/(estimated cfDNA quantity). Applying this detection limit evaluation procedure to the marker negative samples revealed that our ability to detect ctDNA in them was similar to that of the marker positive samples (Supplementary Figure 1-2 and 3A-C).

Tumor and metastasis heterogeneity

We analyzed the initial fresh frozen sample, utilized for mate-pair sequencing, and seven additional tumor FFPE biopsies reflecting different topological locations in the primary tumor by ddPCR (Figure 4). Three SSVs C16D2, C20D1, and C20D2 were found in all biopsies indicating that these two SSVs occurred early in tumor development. Consistent with these observations C16D2, C20D1, and C20D2 also had the most supporting read pairs in the mate-pair analysis and could be detected in the pre-op plasma samples. Analysis of FFPE punch biopsies from the early and late liver lesions revealed that C20D1 and C20D2 were present in both and C16D2 only in the early metastasis (Figure 4D). The sequencing analysis of the late metastasis confirmed that it did not

carry the C16D2 deletion at the RBFOX1 locus seen as the most frequent SSV in both the primary tumor and the early metastasis. Nor did it carry any of the other RBFOX1 deletions seen in the primary tumor (Figure 4A, Supplementary Table 4). A plausible explanation for the observed findings is that early on in tumor development tumorigenic cells carrying only the C20D1 and C20D2 SSVs metastasized to the liver, but lay dormant. Meanwhile, the C16D2 deletion occurred in the primary tumor, in cells harboring the C20D1 and C20D2 SSVs. One of these cells metastasized to the liver, where it expanded to a clinical manifest lesion, which subsequently became completely eliminated by the partial hepatectomy at month 3. Eventually, the dormant C20D1 and C20D2 positive cells started proliferating and formed the late-occurring metastasis.

REFERENCES

1 Korn JM, Kuruvilla FG, McCarroll SA*, et al.* Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nature genetics 2008;**40**:1253-60.

2 Bengtsson H, Wirapati P, Speed TP. A single-array preprocessing method for estimating full-resolution raw copy numbers from all Affymetrix genotyping arrays including GenomeWideSNP 5 & 6. Bioinformatics 2009;**25**:2149-56.

3 Lindgreen S. AdapterRemoval: easy cleaning of next-generation sequencing reads. BMC research notes 2012;**5**:337.

4 Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;**25**:1754-60.

5 Chen K, Wallis JW, McLellan MD*, et al.* BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nature methods 2009;**6**:677-81.

6 Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in bioinformatics 2013;**14**:178-92.

Supplementary Table 1. Summary of samples, assays made, and analyses performed.

Patient ID	Sample ID	Tumor %	SNP	Mate pair	ddPCR	Pipeline 1	Pipeline 2
1	01-blood	N/A*	Yes	No	Yes	Yes	No
2	02-blood	N/A	Yes	No	Yes	Yes	No
4	04-blood	N/A	Yes	No	Yes	Yes	No
5	05-blood	N/A	Yes	Yes	Yes	Yes	Yes
6	06-blood	N/A	Yes	No	Yes	Yes	No
8	08-blood	N/A	Yes	Yes	Yes	Yes	Yes
10	10-blood	N/A	Yes	Yes	Yes	Yes	Yes
15	15-blood	N/A	Yes	No	Yes	Yes	No
16	16-blood	N/A	Yes	Yes	Yes	Yes	Yes
18	18-blood	N/A	Yes	Yes	Yes	Yes	Yes
19	19-blood	N/A	Yes	No	Yes	Yes	No
24	24-blood	N/A	No	Yes	No	No	Yes
28	28-blood	N/A	No	Yes	No	No	Yes
29	29-blood	N/A	No	Yes	No	No	Yes
1	01-tumor	90	Yes	Yes	Yes	Yes	No
2	02-tumor	90	Yes	Yes	Yes	Yes	No
4	04-tumor	80	Yes	Yes	Yes	Yes	No
5	05-tumor	75	Yes	Yes	Yes	Yes	Yes
6	06-tumor	65	Yes	Yes	Yes	Yes	No
8	08-tumor	75	Yes	Yes	Yes	Yes	Yes
10	10-tumor	90	Yes	Yes	Yes	Yes	Yes
15	15-tumor	75	Yes	Yes	Yes	Yes	No
16	16-tumor	70	Yes	Yes	Yes	Yes	Yes
18	18-tumor	90	Yes	Yes	Yes	Yes	Yes
19	19-tumor	75	Yes	Yes	Yes	Yes	No
24	24-tumor	75	No	Yes	No	Yes	Yes
28	28-tumor	65	No	Yes	No	Yes	Yes
29	29-tumor	60	No	Yes	No	Yes	Yes
18	18-metastasis	60	No	Yes	Yes	N/A	Yes
24	24-metastasis	50	No	Yes	No	N/A	Yes
28	28-metastasis	75	No	Yes	No	N/A	Yes
29	29-metastasis	60	No	Yes	No	N/A	Yes

*Not applicable

Supplementary Table 2. Summary of tissue samples and their mate pair characteristics.

Sample ID	MP protocol	Mapped reads	Estimated read length	Useful reads (GB)	Read depth*	Median insert size	Physical read depth
5-blood	Transposon	73,883,169	87	6.4	2.2	2994	37
8-blood	MP library v2	72,990,196	100	7.3	2.5	3122	39
10-blood	MP library v2	82,981,371	100	8.3	2.9	2982	42
16-blood	MP library v2	75,600,526	100	7.6	2.6	2989	39
18-blood	Transposon	95,614,208	83	8.0	2.8	4392	71
24-blood	Transposon	77,825,341	82	6.4	2.2	2529	32
28-blood	Transposon	40,224,075	85	3.4	1.2	1789	12
29-blood	Transposon	83,954,028	84	7.1	2.4	3090	44
1-tumor	MP library v2	21,513,384	100	1.1	0.4	2193	7
2-tumor	MP library v2	67,991,066	100	3.4	1.2	2406	28
4-tumor	MP library v2	33,055,958	100	3.3	1.1	2073	12
5-tumor	MP library v2	42,324,457	100	2.1	0.7	2773	20
6-tumor	MP library v2	30,660,816	100	1.5	0.5	2563	13
8-tumor	MP library v2	48,652,103	100	4.9	1.7	2468	20
10-tumor	MP library v2	39,337,693	100	3.9	1.4	2938	20
15-tumor	MP library v2	37,136,110	100	1.9	0.6	3051	19
16-tumor	MP library v2	37,484,956	100	3.8	1.3	1662	11
18-tumor	MP library v2	25,056,830	100	2.5	0.9	2856	12
19-tumor	MP library v2	75,551,380	100	3.8	1.3	2922	38
24-tumor	MP library v2	64,894,996	100	6.5	2.2	2840	31
28-tumor	MP library v2	29,999,326	100	3.0	1.0	2867	15
29-tumor	MP library v2	18,614,829	100	1.9	0.6	2965	9
18-metastasis	Transposon	95,038,956	87	8.3	2.9	1192	47
24-metastasis	Transposon	92,742,464	84	7.8	2.7	2687	41
28-metastasis	Transposon	102,212,707	84	8.6	3.0	2709	46
29-metastasis	MP library v2	14,900,356	100	1.5	0.5	2902	7

*Read depth equals Useful Reads / 2.897310462

Supplementary Table 3. Primer and TaqMan probe sequences and amplification protocols for PCR validation, Nested PCR, and ddPCR. All probes contain a 6FAM or HEX fluorophore at the 5' end and a black hole quencher-1 (BHQ-1) at the 3' end.

Pt. ID-SSV (Assay)	Sense primer (5'-3') Antisense primer (5'-3') Probe (5'-3'	Amplification protocols
Chr3 (Reference)	CTAGAAGATCTACCTCCAAGAGG CCAGGCTGAAGCTATTCCAG CTCATACATCTGGCATATGGGCTGG	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
gCYC (Reference)	ACATGGGTACTAAGCAACAAAATAAG CACAATTGGAACATCTTTGTTAAAC TTGCAGACAAGGTCCCAAAGACAGCA	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
CPP1 (Purification control)	Reference 1	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
PBC (Lymphocyte contamination control)	Reference 1	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
1-C1I1 (Validation)	GGACCCCAGACAAGTGTGAC GTCCGTGGCTCCACAATTAC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
1-C2D1 (Validation)	CAACTGCCCAAGCAAATACA GGGAGGATTTGAGCTTATCT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
1-C11D1 (Validation)	CTATGCAGGAGATGGGCTTG CTGAGGGTGGGTGGATAAGA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
1-C16D3 (Validation)	TGGTCCCAGTATGGTTAGATGA CCCACCCTCCTTAGGTTTTC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
1-C16d1 (Validation)	CCAGCATCATTCATGGAGTC TTCTGGGGTTTCTGGTGTTT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
2-C1D1 (Validation)	GGCAAGTTTTTGAGGAAAAGG CAGGAGGCAGTCCTGAATTT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
2-C3D1 (Validation)	TCCCTGCTTTAATTTGGAGGT TCCATAGGGTTTGCATATGTCTC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
2-C3D2 (Validation)	TTAGAGGTCATGGCCACATTT GGGTTCAACAGTTCATCAAAGG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
2-C3D3 (Validation)	CACCCACACCAGGCATACA AACCCCAAACTATAAGAATCCTAGC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
2-4D1 (Validation)	GTGTGGGCACATTTGTTCAT GGCATGATGAATGGGGTAAA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
2-C7D1 (Validation)	TTTATGTTAGCTCTATGGTTTTGTAGA AGAAAACATTCCAGGCCAGA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
2-C11D1 (Validation)	GAAGGAGTAGAGTAAACAAGGAAAAGA TTCACTCAGAAAACAGATGCTCA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
2-C13ITX2 (Validation)	GAAGTCCCTTTCTACCCAACCC CAATGGCTAGCCCACCCTCAT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
4-C1D1 (Validation)	TCCCCAACTGTTACTGTACTGC AAGCATTTGACTTTGGCTTG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
4-C2D1 (Validation)	TGTCCTTAAATGTCCTTAAATGC TCTCTGTTAAAACCAGCATTCG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
4-C2D2 (Validation)	GGGTTGAGGATGGAATTTGA CAAAATCATACAAAAGCTCACTCA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
4-C7D1 (Validation)	CCCTAGTCCAGGTGCTTCAG GTCCAGCAAGAGCCCTAGAA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
4-C10D1 (Validation)	ATCCCTTTGAGAGCCAGTCA CCACGGCACTGGGAGTAAAT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
4-C16D1 (Validation)	TTCATAATACAGGAACAAGAGTGTCA GTTCCCCAGGTGAGCTAAAG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
4-C18D1 (Validation)	AGAGTTCCTTTAGGTCTGTTGTGA CCCTGGGCATAGTTGAGAAC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
5-C1D1 (Validation)	CAAGGGCTATCTGGCATCTT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'

5-C1D1 (Validation)	AAGACGTGCAAGCCTTATATTACA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
5-C1D2 (Validation)	GGGTGTGCACTCTCGTCTCT TGCCAGAGTCACAAAAATGG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
5-C1I1 (Validation)	AACAACTGGCACTCAAGAGGA AACTTTTGTCAGCCTTGTGC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
5-C5D1 (Validation)	TGCTTTGATTTTATAGCATGACC AGGAGCTTATGGAAGTCAGGT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
5-C12D1 (Validation)	TGGTCCAGGAAGAGTGAGAC AACAGACACATAGACCAGTGGAAC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
6-C10ITX1 (Validation)	TGTTGTGTTTAAAACCAGTAATTTGA CACCCTAGCTTGCACTGTGA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
6-C20D2 (Validation)	TTCCCCTCCAACAGTGAAAC GCAAAAGGTTTTGAAAATGG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
6-20D3 (Validation)	CTAAGCACAGAAGGTACTAGAGAACA GAAACTGCCTCGCAAATCAT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
6-C20I1 (Validation)	TGTCAGGGTCAGGATTTGAA GGCTGCTCCACAGACACTG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
6-C20ITX1 (Validation)	TTCTGACTGGGGATTTTTCTG AATCCTCAATCCGACTTTGG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
8-C1D1 (Validation)	CCATTCCACACCCACATACT GTTGGCGGAGTTCAGAATTA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
8-C3D1 (Validation)	GTAAGATCGGGAGGGAAGAA CAGGAGTGGCTCTTCTCAAA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
8-C7D1 (Validation)	TAGGGGTGCTCCACCTTCAG CTGAAACCTCTGCTTTGGAACG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
8-C15D1 (Validation)	GTAAGCGACCTCAGCGTTT GCTTGTCAATCTTAGACACTTGAAC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
8-C16D1 (Validation)	ATGGATTTGCCTATGTCCAA GCTGGCATTTCAGTGTTCTC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
8-C16D2 (Validation)	CCTTAAGTTGCGCCTGTTTA GTGAGCATCCAAGGTAGGAA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
8-C16D3 (Validation)	TGACCAGGTGTGGATGGAGA AGTACGGCAGTGCAAACACA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
10-C3D1 (Validation)	AAGTTTTTGTCTTGGCTTGG GCTGGGTACTCCTCTCCTGT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
10-C6D1 (Validation)	GTGTCAGACTGAAGAGGTATGG GCCATAGTTAGCCAATCCCTA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
10-C7D1 (Validation)	AAACTCGCTCCAGATCATCC TGTCAGTTGTACCTGCAGCT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
10-C7D2 (Validation)	CCCAGAACGGTGAGTAAATG CTGCAAAATCATGCCAAAA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
10-C9D1 (Validation)	TGGCTTTCAATGGGATCATA AGAAACCGGTTGGAGGTAAC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
15-C12D1 (Validation)	GGGAGGGTGAGAAAGAGGAG TCATTTCCCTGCTCCAAAAC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
15-C12I1 (Validation)	ATCTGGCTTGTAAAACTGGGTG CAATGGGGGAAACTGGAAAG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
15-C12D2 (Validation)	TTTAATCAGTTTTTCTTCCATCCT AATGCCTTGGATGAGCTGAG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
15-C12I2 (Validation)	GTCTTGTCCCCTTCGCAGAC CAGCTTCCATTGTCCCTCTC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
15-C14D1 (Validation)	CCTTCACTGTCGCTGAGGAT TGATCTTAAAGCTACAGAGTGACTGG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
16-C6D3 (Validation)	GGAATCCGACTACAGGGTTT TGCCATTAGACTGGACTGGT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
16-C6D1 (Validation)	TGAACCAAGAGCAAACAATC CTTTGCCCATGTTCTCTGTC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
16-C14D1 (Validation)	TTGAGTAGCTCTCCCACCTG CTATGGCCCCCTATTCAACT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'

16-C16D1 (Validation)	AGCCCTCCTCTGCAATAGTT CACACCTGACACAGCACATT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
18-C3D1 (Validation)	TCCTGCTGAATTACCTATTCG GAAAAGGAAACAAATGCACAG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
18-C16D1 (Validation)	CAAAGGCAAGGACAGAAATG CACACATCGAGATGCTCTGT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
18-C16D1 (Validation)	CAGGAGACAAAGCTTACTGCC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
18-C16D2 (Validation)	CCCATCACTGCATCATATTG ATTTACTTTGGGGCTCATGG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
18-C16D3 (Validation)	TTGCAAGCTATTCAGTCAACTC TCCTTTCTTTCCTTCAGTGG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
18-C20D1 (Validation)	CAATAAACCCCACAAACAGC GGCCACCTATTGACATCATC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
18-C20D2 (Validation)	CCCCTCTGCACTTAGTGTGT TATTCCTCCATGCCACTGAT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
19-C5D1 (Validation)	AAACGAAAAGGAAGAAAAATGA CTGTCCCAGTTTGGGTTTTC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
19-C8I1 (Validation)	CTAACTCAAGACACAAGGCAC TCCTTTCATAGGGATACAACATGG	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
19-C12D1 (Validation)	AACAGCCCTACCTTCCCACT CCACTCTACATTCCACAGATGC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
19-C12D2 (Validation)	AACAGCCCTACCTTCCCACT CCACTCTACATTCCACAGATGC	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
19-C13T(13:8)1 (Validation)	CTCTGCCTCACACCCAGAA CCGAAATGTCCAATAGCAAA	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
19-C13T(13:8)2 (Validation)	GAAATGAATAAGCTGGTTTTGCT CTCTGGGAGGCTAAACTTTTGT	95°C 15', (95°C 30'', 58°C 30'', 72°C 2') x40 & 72°C 5'
1-C1I1 (ddPCR)	GCAGGAGGTACACAGGAAGTAT TTGCCACCTTCCATAGATTT TGTGAGGCAAGAGCAGTCTGG	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
1-C2D1 (ddPCR)	TTTGTTTCTTAATTAATGCCTTTGA TCAGGGGATTTGTTTGAAGTC CCTACTGGTTGCAGCAGCTTCA	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
1-C11D1 (ddPCR)	GACCCTACCTATGGTTGTTGAGAGT AGCTGGGTGTATTGTAATTCCCACA AATGGCCCACGGGGTCTGCTG	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
1-C16D3 (ddPCR)	TGCTGCACATTTACACCGTTC TTTCTCCCAACTTCAATCATATACCA TAGGCTTTCGCATTTTCACCTTACCACC	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
1-C16d1 (ddPCR)	GGCTCATGTGGTTCAGGAAG CGCATGGAACCAAGATGAAT GAATGCAGGGAGTTATTAAAAACCTCA	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
2-C1D1 (ddPCR)	TTGGCTTGGAGATATCTGAAAAGTGT TGCACTAGAAATTCAGTCATGCCT AATGAGGGTGGAAATAAGCTTGCCTCT	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
2-C3D1 (ddPCR)	GATCTAGGGGTTAAACTCTACTGAA TGCAGCCCTTTGTATATTGGATG TATGGTTGCTTTTGGTGGAAACCCCCTT	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
2-4D1 (ddPCR)	CATAACTTTTTCAACAAGCATCCA GGGCACATTTGTTCATAATATTTC TCCTTTAGGCCCAATATCTGAAAGATAA	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
2-C11D1 (ddPCR)	ATTTGGCACTATCAGCAAAGTATTA AGTTCTGGGTACAATTCTAGAGC ACTGAAATTAAGGCAGAGAGACCTGGA	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
2-C13ITX2 (ddPCR)	TTCACTCTCATTGCTGAGTGGT AGAGAAGCCCACACAGGTTG TTGCCGACTCACCTGCTTCT	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'

4-C1D1d (ddPCR)	TCCCCAACTGTTACTGTACTG ATCTATTCAAGGCATATATAGCTACA TTGCTGTACTCAACTCTGAAGGACT	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
4-C2D1d (ddPCR)	TTTGTTCTGTTTATGGAGATGAAG AGCTGAACCATATAGGCTG ATTCCAAGAAAACCAAAGTTGACCCA	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
4-C7D1d (ddPCR)	GCCTTCCTAAATCCTCATGTGT CTATATGAGAAAGGGAAAATTTGTGT ACCTGGTCTGCCACAATTCCAC	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
4-C10D1d (ddPCR)	TATGTTATACAAAATTGTCCCACGAG TGATTTCCTGTCCAAGGTGC ACACTGAAGTATGCCCACGGTT	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
5-C1D2 (ddPCR)	TTGGTCTTCCTGCTCAATCTC GGAGTGCAGTGGTGCTGAG TTTTGACATACCCTATTCTCCTGC	95°C 10', (95°C 30'', 57°C 30'') x45 & 98°C 10'
5-C1I1 (ddPCR)	CTTCGGGACCTTGAACTTTG TTTGGGTGTATACCCGTTCC TCCATATTCTTGGAACTGTGTACCCA	95°C 10', (95°C 30'', 57°C 30'') x45 & 98°C 10'
5-C5D1 (ddPCR)	TGACCTGTATGAGTAAGCATATAAAAT TGGTTTTGTGTTATAATCTTGTTCTC TTAGAGCCTACTAAAATGATCAAGATCTC	95°C 10', (95°C 30'', 61°C 30'') x45 & 98°C 10'
5-C12D1 (ddPCR)	GCCACACAGCTGATTTAGAGG ACCCCACAAGAAAACCAAAG TCTTCTGGCCATTACTCTTAAGCA	95°C 10', (95°C 30'', 57°C 30'') x45 & 98°C 10'
6-C10ITX1 (ddPCR)	CCGGGTTCTTTCCCTTTCT TTTATACCCCAGACAGTTCTCCA ACCTCCTGTGATCCTGGCCC	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
6-C20D2 (ddPCR)	AAGGATTGTGTGTATCTGTGATGTC CTCTGTCCTGCCTCCCTCTA AAATCATACATTAGGGGATTTTGTTC	95°C 10', (95°C 30'', 56°C 30'') x45 & 98°C 10'
6-20D3 (ddPCR)	TTCCTCCATGAGCCCTCTC CACTGTTCATCCATAGCAGCTC CCACCACACAATGCCAAGCA	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
6-C20I1 (ddPCR)	ACACCGCCAAGGTGAATTT TCCGGTCTATCACCAGCTTC CACCTAGAGGTCCCGGCAGG	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
8-C1D1d (ddPCR)	CGCAACATGAATGAATGTCAGA AGCCAAGCCAGAGAGAGG AGTTCCAATCCTTCCCAGTCAGCTGCA	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
8-C3D1d (ddPCR)	TGTCTTCCCTGAGTGACCAT ACGCTATAAATTGGCCTTTG CAATGAAGCCAGAGACCATGCATG	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
8-C7D1d (ddPCR)	TCACAGCTTCAACAAGTTTGC TCAAGCTCTAACAGTCTTCAAACTAA TGGAACAACTGTAAGAGATGTGGTTTG	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
8-C15D1d (ddPCR)	CCCTGTGGGCTGCAAAGATT ACAGAACCAGGTGAGAGCTG CCACCCTGGTGCTGTCCTTGCAGCTCA	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
8-C16D2d (ddPCR)	AGAGCTTACAAGAGCACCT AGTATTTATAGTCCTTCAGTACCTCG CTCTCTGAGCACTCCCCGTGT	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
10-C3D1N (Nested)	TATGCCAGAGATTGCAAATGT CCAGCAAGCTATCACGTAGG	95°C 15', (95°C 30'', 58°C 30'', 72°C 45'') x12 & 72°C 5'
10-C3D1d (ddPCR)	GCCAGAGATTGCAAATGTTT TGGAAAAACTGAGTCACATGAA CATAAGTTTAGTGTTCCTGGCTCGGGA	95°C 10', (95°C 30'', 58°C 30'') x45 & 98°C 10'
10-C6D1N (Nested)	GGGACTTTGAAATTCTGAGATG	95°C 15', (95°C 30'', 58°C 30'', 72°C 45'') x12 & 72°C

10-C6D1N (Nested)	GAGGTGTTTTAATTAACAATCTTGAG	5’
10-C6D1d (ddPCR)	TCCATCAGGAAATTTTTGG TGTGCTATGTTTTTGAAGCA CTTGGAGAGATTTGTTTACACATGTGTAGG	95°C 10’, (95°C 30‘‘, 58°C 30‘‘) x45 & 98°C 10’
10-C7D2N (Nested)	AAGTTTTGTTGGCTTCGAGA CAATACAGGAGCACCCAGAT	95°C 15’, (95°C 30‘‘, 58°C 30‘‘, 72°C 45‘‘) x12 & 72°C 5’
10-C7D2d (ddPCR)	CCAGAATCTCCAGAGCCA CCACACATTAATAATGGGAGACT CTCCATTCCTTTAGAACCACTACAGTGGG	95°C 10’, (95°C 30‘‘, 58°C 30‘‘) x45 & 98°C 10’
10-C9D1N (Nested)	CCATGTTTACACCACGGAAC AGAAACCGGTTGGAGGTAAC	95°C 15’, (95°C 30‘‘, 58°C 30‘‘, 72°C 45‘‘) x12 & 72°C 5’
10-C9D1d (ddPCR)	TCAATGCTCCACTATTGGTAAA CGGTTGGAGGTAACTGAATC CTATTTCTAAGTGAACTATCAGGAGAAC	95°C 10’, (95°C 30‘‘, 58°C 30‘‘) x45 & 98°C 10’
15-C12D1 (ddPCR)	TCTATGTTCAGAGAGGCCTGATTT AGCAAATAAGCTGTCACTGTTGAG AGCCACCTGCCTGGGGTCTTCA	95°C 10’, (95°C 30‘‘, 59°C 30‘‘) x45 & 98°C 10’
15-C12I2 (ddPCR)	GGTTGACCTGTTGTCCAAGT AAGGCATCTAAACAGGAAAAGA AGAACATTCGTCCCTTCTCTTCTCC	95°C 10’, (95°C 30‘‘, 59°C 30‘‘) x45 & 98°C 10’
16-C6D3N (Nested)	AGGCTAAAACTCTTGGAATTTTCT TTCTTTCCAATGTTAAATCTTGC	95°C 15’, (95°C 30‘‘, 58°C 30‘‘, 72°C 45‘‘) x12 & 72°C 5’
16-C6D3d (ddPCR)	GAGTCCCTTTTGATCACACC GCCAGAAAATTCATCCAGTT AATTTATTGGATTTGCTTGGGAAGAGTGG	95°C 10’, (95°C 30‘‘, 58°C 30‘‘) x45 & 98°C 10’
16-C14D1N (Nested)	AAACCCCTTTTCTCCATGAC TGCATAGAAAGTTTGATTTGCTT	95°C 15’, (95°C 30‘‘, 58°C 30‘‘, 72°C 45‘‘) x12 & 72°C 5’
16-C14D1d (ddPCR)	CCTTTTCTCCATGACCTTTG TCATGAGTTCCATTGATTGC CACAGTATTCTCTCTATTGCAGTGGTGCT	95°C 10’, (95°C 30‘‘, 58°C 30‘‘) x45 & 98°C 10’
16-C16D1N (Nested)	AGAGGCTGATCAAAGGGAAT CCCACTGAGATGATTTTTGC	95°C 15’, (95°C 30‘‘, 58°C 30‘‘, 72°C 45‘‘) x12 & 72°C 5’
16-C16D1d (ddPCR)	GGGATTGCAGGACTCTTTC GAATGATCACAGCCTTCAAA AGGGGAAGGAGGCAGAGTGCTTA	95°C 10’, (95°C 30‘‘, 58°C 30‘‘) x45 & 98°C 10’
18-C3D1d (ddPCR)	CTGCTTTACACTCAGTTGATCTGT GAGTCACCAGAGACCTTGTCACA TTTCCCCTGCCTAGCCTTGCCT	95°C 10’, (95°C 30‘‘, 59°C 30‘‘) x45 & 98°C 10’
18-C16D1d (ddPCR)	CCTCGCTCAAGAAGTCTGT TTCCCTGAAACTCCCACAACA TCCTACGGGGAGAGACTCAATGC	95°C 10’, (95°C 30‘‘, 59°C 30‘‘) x45 & 98°C 10’
18-C16D2d (ddPCR)	ATGGAATGCTGTATTCATACATCTGA TTACTTTGGGGCTCATGGACTTG CCAGATGTGGATACCCTGGAAAGTGG	95°C 10’, (95°C 30‘‘, 59°C 30‘‘) x45 & 98°C 10’
18-C16D3d (ddPCR)	ACTGTACCAAGCAACTACATT ATTCAAACAGACTGAGTGTTGAC ATGCCTAAGACTTTAAGAGGAAAAGCT	95°C 10’, (95°C 30‘‘, 59°C 30‘‘) x45 & 98°C 10’
18-C20D1d (ddPCR)	ACAACCCCAGGAATTAGAAGG TCTGTGAGTCTGGTTAGCACAA AAGACGGCATGAGTGAGTTCCCT	95°C 10’, (95°C 30‘‘, 59°C 30‘‘) x45 & 98°C 10’
18-C20D2d (ddPCR)	TTCATTACCCTGAGCTTTGGTT TCTTCAGAGATTCCCCCTTCC TAATCTGGGCAGCCTCCAGTTCACA	95°C 10’, (95°C 30‘‘, 59°C 30‘‘) x45 & 98°C 10’
19-C5D1 (ddPCR)	CAGAAGGAATCATAATCAATTTTCTA CAAAGTGATATCCTCAAACAGCTA TGAAAGGAAGGTCTAGCTTAATCATT ACCTTGGGCAAGAGAGGTTA	95°C 10’, (95°C 30‘‘, 59°C 30‘‘) x45 & 98°C 10’

19-C8I1 (ddPCR)	AAAATAGGTCATCAAAAGGAAAAA CACTGAATCTTTGTGGCCTCACC	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
19-C12D1 (ddPCR)	TGTGCTCCAAATCAAAATGC GTCATTGCCACTGGTCTCCT TCCAGATCACTAGCTCTGAACATTGA	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
19-C12D2 (ddPCR)	ACTGATACTGGTTAAGAGGATCTTG TTCCATTTTAATGGGAACAATG TCCCAAATGGAATCTATAAAGGGA	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
19-C13T(13:8)1 (ddPCR)	GGCAATTTCCCCTGTATCC AATAGTGTTTCTCTGGGGATGA TGAGAATTAAAAATTGACCTTCTCCA	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'
19-C13T(13:8)2 (ddPCR)	TTACAACCATTGTTGAGGGTCT TTGTTCAAGCCAATGTGTTTTC CTGTCCAGGGAATAAAAATGTTTGA	95°C 10', (95°C 30'', 59°C 30'') x45 & 98°C 10'

1Pallisgaard N AR, Spindler KG, Jakobsen A. European Journal of Cancer.Vol 49:suppl.4, Abstract MC13-088, 2013.

Supplementary Table 4. Somatic rearrangements identified by pipeline 1 and 2.

Pt.	Pipeline	Chr.	Location 1 (Hg19)	Chr.	Location 2 (Hg19)	Size(bp)	Class	Reads	Amplitude (SNP)[#]	Validated by Sanger sequencing	Plasma SSV analysis by ddPCR	Preserved from tumor to metastases
1	1	1	21,800,307	1	64,580,322	42,780,015	INV	4	-0.462	C1I1	Negative	N/A**
1	1	2	190,380,577	2	190,403,932	23,355	DEL	5	-0.345	C2D1	Negative	N/A
1	1	4	99,212,690	4	99,282,549	69,859	DEL	2	-1.338		ND*	N/A
1	1	5	59,509,055	5	59,643,263	134,208	DEL	2	-0.403		ND	N/A
1	1	7	90,068,750	7	90,270,995	202,245	DEL	4	-0.559		ND	N/A
1	1	11	84,971,740	11	85,313,110	341,370	DEL	3	-0.352	C11D1	Negative	N/A
1	1	16	6,344,378	16	6,940,220	595,842	DEL	4	-0.523	C16D3	Negative	N/A
1	1	16	6,459,598	16	7,211,090	751,492	DEL	4	-1.165		ND	N/A
1	1	16	7,418,474	16	7,496,667	78,193	DEL	3	-0.536	C16D1	Negative	N/A
1	1	18	57,549,267	18	57,639,302	90,035	DEL	3	-1.096		ND	N/A
1	1	20	14,719,275	20	15,258,734	539,459	DEL	4	-0.405		ND	N/A
1	1	20	14,922,380	20	15,022,536	100,156	DEL	3	-0.42		ND	N/A
1	1	20	14,969,501	20	15,195,944	226,443	DEL	2	-0.405		ND	N/A
2	1	1	192,775,425	1	192,821,726	46,301	DEL	5	-1.04	C1D1	Negative	N/A
2	1	3	60,138,611	3	60,196,513	57,902	DEL	4	-0.44	C3D1	Negative	N/A
2	1	3	60,290,731	3	60,393,493	102,762	DEL	9	-0.49	C3D2	ND	N/A
2	1	3	60,484,029	3	60,661,659	177,630	DEL	7	-0.48		ND	N/A
2	1	3	174,764,067	3	174,865,023	100,956	DEL	7	-0.86	C3D3	ND	N/A
2	1	3	174,865,300	3	174,976,187	110,887	DEL	7	-0.36		ND	N/A
2	1	4	91,766,206	4	91,818,287	52,081	DEL	7	-0.34		ND	N/A
2	1	4	93,177,692	4	93,330,435	152,743	DEL	4	-0.37		ND	N/A
2	1	4	93,400,494	4	93,916,277	515,783	DEL	6	-0.30		ND	N/A
2	1	4	93,586,033	4	93,698,710	112,677	DEL	4	-0.96		ND	N/A
2	1	4	106,175,087	4	106,525,867	350,780	DEL	11	-0.37	C4D1	Positive	N/A
2	1	6	19,185,392	6	40,968,849	21,783,457	INV	8	-0.42		ND	N/A
2	1	6	91,712,449	6	91,856,061	143,612	DEL	6	-0.4		ND	N/A
2	1	6	163,050,354	6	163,162,927	112,573	DEL	6	-0.92		ND	N/A
2	1	7	145,992,207	7	146,306,320	314,113	DEL	5	-0.4		ND	N/A
2	1	7	146,323,715	7	146,395,251	71,536	DEL	3	-0.397	C7D1	ND	N/A
2	1	7	146,407,661	7	146,618,400	210,739	DEL	5	-0.855		ND	N/A
2	1	7	146,504,749	7	146,691,204	186,455	DEL	3	-0.414		ND	N/A

2	1	7	146,852,920	7	146,945,474	92,554	DEL	4	-0.987		ND	N/A
2	1	7	147,057,615	7	147,206,451	148,836	DEL	7	-0.469		ND	N/A
2	1	11	80,646,575	11	80,977,206	330,631	DEL	3	-0.4	C11D1	Negative	N/A
2	1	13	27,530,543	13	27,537,906	7,363	ITX	91	1.238	C13ITX2	Positive	N/A
2	1	20	14,808,980	20	14,848,561	39,581	DEL	2	-0.384		ND	N/A
2	1	20	14,821,503	20	15,172,003	350,500	DEL	3	-0.6		ND	N/A
2	1	20	14,966,177	20	15,201,351	235,174	DEL	2	-0.3		ND	N/A
4	1	1	106,503,709	1	118,449,060	11,945,351	DEL	4	-0.127	C1D1	Positive	N/A
4	1	2	110,109,317	2	110,251,577	142,260	DEL	5	-0.339		ND	N/A
4	1	2	110,874,389	2	111,037,146	162,757	DEL	5	-0.339		ND	N/A
4	1	2	114,553,679	2	114,619,945	66,266	DEL	3	-0.246		ND	N/A
4	1	2	118,051,890	2	122,662,329	4,610,439	DEL	5	-0.378		ND	N/A
4	1	2	136,631,841	2	137,959,724	1,327,883	DEL	2	-0.163	C2D2	ND	N/A
4	1	2	143,775,511	2	149,312,728	5,537,217	DEL	2	-0.374	C2D1	Positive	N/A
4	1	5	146,440,990	5	146,878,844	437,854	DEL	2	-0.269		ND	N/A
4	1	6	381,150	6	625,549	244,399	DEL	3	-0.289		ND	N/A
4	1	6	1,128,938	6	2,163,475	1,034,537	DEL	3	-0.33		ND	N/A
4	1	6	2,237,480	6	5,325,143	3,087,663	DEL	2	-0.318		ND	N/A
4	1	6	5,587,549	6	5,925,022	337,473	DEL	2	-0.354		ND	N/A
4	1	6	6,882,525	6	9,158,626	2,276,101	DEL	2	-0.342		ND	N/A
4	1	6	20,781,286	6	20,901,416	120,130	DEL	2	-0.357		ND	N/A
4	1	6	21,010,568	6	21,347,173	336,605	DEL	2	-0.308		ND	N/A
4	1	6	21,421,445	6	21,462,849	41,404	DEL	1	-0.378		ND	N/A
4	1	6	22,015,481	6	22,074,193	58,712	DEL	2	-0.322		ND	N/A
4	1	6	65,846,617	6	65,876,997	30,380	DEL	1	-0.253		ND	N/A
4	1	7	147,155,588	7	147,805,585	649,997	DEL	2	-0.122	C7D1	Positive	N/A
4	1	8	60,726,644	8	82,048,061	21,321,417	AMP	3	0.237		ND	N/A
4	1	8	90,312,120	8	91,641,447	1,329,327	AMP	4	0.246		ND	N/A
4	1	10	134,652,722	10	135,252,996	600,274	DEL	1	-0.355		ND	N/A
4	1	10	28,421,452	10	28,870,756	449,304	DEL	2	-0.277	C10D1	Positive	N/A
4	1	14	95,780,574	14	96,734,831	954,257	DEL	2	-0.388		ND	N/A
4	1	14	96,736,016	14	96,797,473	61,457	DEL	2	-0.282		ND	N/A
4	1	14	96,980,052	14	97,124,096	144,044	DEL	4	-0.318		ND	N/A
4	1	14	97,151,378	14	97,226,233	74,855	DEL	1	-0.443		ND	N/A
4	1	16	6,653,132	16	6,752,046	98,914	DEL	2	-0.145	C16D1	ND	N/A

4	1	17	11,590,136	17	11,997,692	407,556	DEL	1	-0.683		ND	N/A
4	1	18	2,763,210	18	4,038,123	1,274,913	DEL	2	-0.404	C18D1	ND	N/A
5	2	1	26,213,965	1	26,579,546	364,025	DEL	2		C1D1	ND	N/A
5	2	1	36,670,445	1	55,589,198	18,917,576	DEL	2		C1D2	Negative	N/A
5	2	1	55,429,415	1	59,046,906	3,616,905	INV	2		C1I1	Negative	N/A
5	2	5	119,499,524	5	119,549,363	52,620	DEL	3		C5D1	Negative	N/A
5	2	12	86,881,455	12	86,970,252	91,383	DEL	2		C12D1	Negative	N/A
6	1	10	5,382,443	10	5,782,602	400,159	ITX	3	0.208	C10ITX	Positive	N/A
6	1	20	14,677,101	20	15,070,699	393,598	DEL	3	-0.125	C20D2	Negative	N/A
6	1	20	29,809,848	20	54,457,627	24,647,779	DEL	2	0.312		ND	N/A
6	1	20	30,326,644	20	47,981,511	17,654,867	ITX	2	0.493		ND	N/A
6	1	20	39,815,996	20	57,603,751	17,787,755	ITX	4	0.432		ND	N/A
6	1	20	42,817,826	20	61,090,236	18,272,410	DEL	4	0.656	C20D3	Positive	N/A
6	1	20	42,828,701	20	56,544,922	13,716,221	ITX	7	0.656		ND	N/A
6	1	20	48,277,025	20	61,058,932	12,781,907	INV	2	0.39		ND	N/A
6	1	20	51,106,900	20	61,090,236	9,983,336	INV	5	0.473	C20I1	Positive	N/A
6	1	20	51,590,161	20	60,118,999	8,528,838	ITX	7	0.355	C20ITX1	ND	N/A
6	1	20	52,281,876	20	61,003,433	8,721,557	DEL	2	0.57		ND	N/A
6	1	20	56,557,247	20	61,058,932	4,501,685	INV	4	0.358		ND	N/A
6	1	20	60,114,377	20	61,049,656	935,279	INV	3	0.357		ND	N/A
8	1	7	11,555,334	7	12,509,227	953,893	DEL	7	-0.299	C7D1	Positive	N/A
8	2	1	245,427,132	1	245,614,918	187,786	DEL	2			ND	N/A
8	2	10	5,618,677	10	5,642,936	24,259	ITX	5			ND	N/A
8	2	10	27,378,773	10	75,270,851	47,892,078	ITX	2			ND	N/A
8	2	15	75,624,637	15	78,238,153	2,613,516	INV	2			ND	N/A
8	2	16	5,436,326	16	5,906,076	469,750	DEL	3			ND	N/A
8	1+2	1	5,686,842	1	37,310,730	31,623,888	DEL	3	-0.213	C1D1	Positive	N/A
8	1+2	3	59,848,514	3	60,212,346	363,832	DEL	3	-0.294	C3D1	Negative	N/A
8	1+2	15	80,743,250	15	86,355,604	5,612,354	DEL	5	-0.216	C15D1	Negative	N/A
8	1+2	15	81,234,423	15	86,360,900	5,126,477	ITX	3	-0.218		ND	N/A
8	1+2	15	85,094,942	15	90,579,064	5,484,122	INV	2	-0.216		ND	N/A
8	1+2	16	6,298,171	16	6,455,087	156,916	DEL	3	-0.301	C16D1	ND	N/A
8	1+2	16	6,457,503	16	6,889,515	432,012	DEL	6	-0.301	C16D2	Negative	N/A
8	2	16	5,436,326	16	5,906,076	469,750	DEL	3		C16D3	ND	N/A
10	1	7	144,859,153	7	145,710,629	851,476	DEL	1	-0.523	C7D2	Positive	N/A

10	2	2	241,552,388	2	241,809,817	257,429	DEL	4			ND	N/A
10	2	3	44,347,960	3	44,370,945	22,985	DEL	11		C3D1	Positive	N/A
10	2	5	146,024,823	5	146,065,960	41,137	DEL	2			ND	N/A
10	2	6	6,147,180	6	6,851,738	704,558	DEL	3			ND	N/A
10	2	7	63,319,538	7	73,874,442	10,554,904	DEL	3		C7D1	ND	N/A
10	2	8	60,013,119	8	61,680,052	1,666,933	INV	11			ND	N/A
10	2	8	60,013,119	8	61,689,902	1,676,783	INV	9			ND	N/A
10	2	16	70,155,648	16	74,394,148	4,238,500	INV	2			ND	N/A
10	1+2	1	6,225,150	1	36,045,932	29,820,782	DEL	4	-0.49		ND	N/A
10	1+2	3	99,995,570	3	118,411,438	18,415,868	INV	2	0.228		ND	N/A
10	1+2	3	99,995,570	3	118,414,770	18,419,200	DEL	4	0.228		ND	N/A
10	1+2	4	171,659,948	4	174,060,373	2,400,425	DEL	2	-0.49		ND	N/A
10	1+2	6	114,209,636	6	114,274,860	65,224	DEL	3	-1.195	C6D1	Negative	N/A
10	1+2	9	90,381,766	9	91,411,294	1,029,528	DEL	5	-0.509	C9D1	Positive	N/A
15	1	5	159,084,546	5	159,320,294	235,748	ITX	3	0.33		ND	N/A
15	1	11	56,624,833	11	65,908,165	9,283,332	DEL	7	-0.265		ND	N/A
15	1	11	66,442,472	19	35,440,503	-31,001,969	CTX	4	-0.304		ND	N/A
15	1	12	2,581,688	12	3,521,403	939,715	DEL	26	-0.359	C12D2	ND	N/A
15	1	12	2,741,301	12	15,090,194	12,348,893	INV	6	-0.285	C12I2	Positive	N/A
15	1	12	3,526,678	12	4,973,443	1,448,586	DEL	20	0.997	C12D1	Positive	N/A
15	1	12	4,055,544	12	14,759,164	10,703,620	INV	30	0.938	C12I1	ND	N/A
15	1	14	31,727,802	14	32,312,933	585,131	DEL	5	-0.283	C14D1	ND	N/A
15	1	15	89,850,225	15	90,165,845	315,620	ITX	3	0.387		ND	N/A
15	1	17	28,555,679	17	34,456,119	5,900,440	DEL	7	-0.274		ND	N/A
15	1	17	34,257,406	17	34,935,790	678,384	INV	6	0.757		ND	N/A
15	1	17	39,314,104	17	40,333,669	1,019,565	INV	3	-0.274		ND	N/A
16	1	14	42,004,790	14	49,754,826	7,750,036	DEL	1	-0.271	C14D1	Positive	N/A
16	1	16	7,035,042	16	7,146,431	111,389	DEL	1	-0.272		ND	N/A
16	2	1	77,695,769	1	77,794,867	99,098	DEL	2			ND	N/A
16	2	1	95,130,572	1	95,314,672	184,100	DEL	3			ND	N/A
16	2	4	26,324,727	4	168,420,548	142,095,821	INV	2			ND	N/A
16	2	6	121,461,719	6	121,630,100	168,381	DEL	4		C6D1	ND	N/A
16	2	6	5,081,051	6	5,523,350	442,299	DEL	2		C6D3	Positive	N/A
16	2	9	136,182,707	9	138,410,139	2,227,432	DEL	2			ND	N/A
16	2	10	77,628,406	10	77,665,811	37,405	DEL	2			ND	N/A

16	2	12	100,433,893	12	100,607,683	173,790	DEL	2			ND	N/A
16	2	14	53,904,951	14	54,129,135	224,184	DEL	2			ND	N/A
16	2	15	64,964,884	15	67,197,464	2,232,580	INV	5			ND	N/A
16	2	15	67,193,914	15	67,632,363	438,449	INV	2			ND	N/A
16	2	20	34,263,207	20	34,290,057	26,850	ITX	2			ND	N/A
16	1+2	16	6,372,175	16	6,685,322	313,147	DEL	2	-0.271		ND	N/A
16	1+2	16	6,688,080	16	7,034,941	346,861	DEL	2	-0.599	C16D1	Positive	N/A
18	1	20	15,013,498	20	15,146,115	130,521	DEL	1	-0.31		ND	No
18	2	3	49,444,117	12	86,120,327	N/A	CTX	2			ND	No
18	2	3	59,811,840	3	59,905,790	93,950	DEL	2			ND	No
18	2	3	175,396,006	10	80,566,192	N/A	CTX	5			ND	Yes
18	2	7	110,429,796	7	110,481,092	51,296	DEL	3			ND	No
18	2	11	61,841,732	14	81,786,446	N/A	CTX	2			ND	Yes
18	2	16	47,991,864	16	69,162,992	21,171,128	ITX	2			ND	No
18	2	16	85,204,195	16	86,503,113	1,298,918	INV	4			ND	Yes
18	2	17	37,792,424	17	39,773,331	1,980,907	ITX	2			ND	No
18	2	20	29,535,048	20	30,535,784	1,000,736	DEL	2			ND	No
18	2	20	32,257,728	20	32,503,263	245,535	DEL	2			ND	No
18	2	21	30,486,042	21	30,546,052	60,010	ITX	4			ND	No
18	1+2	3	25,788,478	3	29,942,146	4,153,668	DEL	2	-0.241	C3D1	Negative	No
18	1+2	9	4,376,242	9	7,805,854	3,429,612	INV	4	-0.743		ND	Yes
18	1+2	9	4,789,944	9	9,665,095	4,875,151	INV	6	-0.804		ND	Yes
18	1+2	16	6,719,893	16	6,802,435	82,542	DEL	3	-0.22	C16D1	Negative	No
18	1+2	16	6,786,299	16	6,942,699	156,400	DEL	6	-0.201	C16D2	Positive	Yes[$]
18	1+2	16	6,806,586	16	6,865,912	59,326	DEL	3	-0.894	C16D3	Negative	No
18	1+2	20	14,842,321	20	15,326,591	484,270	DEL	3	-0.802	C20D1	Positive	Yes
18	1+2	20	15,254,358	20	15,363,390	109,032	DEL	4	-1.27	C20D2	Positive	Yes
19	1	5	93,036,151	5	116,904,266	23,868,115	DEL	7	-0.308	C5D1	Negative	ND
19	1	8	39,232,907	8	39,387,407	154,500	DEL	5	0.39		ND	ND
19	1	8	115,867,302	13	82,256,869	-33,610,433	CTX	10	0.353	C13T(13:8)2	Positive	ND
19	1	8	124,152,277	13	83,045,953	-41,106,324	CTX	13	0.629	C13T(13:8)1	Negative	ND
19	1	8	136,659,089	8	137,983,171	1,324,082	INV	14	0.361	C8I1	Negative	ND
19	1	12	750,451	12	1,673,364	922,913	DEL	26	0.308	C12D1	Positive	ND
19	1	12	750,451	12	1,673,364	922,913	DEL	26	0.308	C12D2	Positive	ND
24	2	1	168,188,764	1	182,274,074	14,085,310	INV	3			N/A**	Yes

24	2	3	193,862,627	3	193,985,240	122,613	ITX	2		N/A	Yes
24	2	4	13,648,443	4	15,503,473	1,855,030	DEL	3		N/A	Yes
24	2	8	28,435,150	17	79,759,475	N/A	CTX	3		N/A	Yes
24	2	9	130,557,710	9	130,714,806	157,096	DEL	3		N/A	No
24	2	13	86,973,917	13	87,951,805	977,888	INV	6		N/A	Yes
24	2	16	6,621,165	16	6,672,258	51,093	DEL	2		N/A	Yes
24	2	17	19,427,192	17	19,790,533	363,341	INV	2		N/A	Yes
24	2	17	21,074,252	17	21,292,903	218,651	DEL	3		N/A	Yes
24	2	20	14,850,153	20	15,072,373	222,220	DEL	2		N/A	Yes
24	2	23	41,184,610	23	78,676,440	37,491,830	INV	4		N/A	Yes
28	2	1	241,431,223	1	242,359,929	928,706	ITX	2		N/A	Yes
28	2	2	98,384,391	2	99,501,960	1,117,569	DEL	3		N/A	Yes
28	2	4	170,281,585	14	52,664,801	N/A	CTX	2		N/A	Yes
28	2	4	91,139,559	4	92,139,761	1,000,202	DEL	4		N/A	No
28	2	5	112,175,701	5	112,240,950	65,249	DEL	3		N/A	Yes
28	2	6	77,436,862	6	77,461,604	24,742	DEL	2		N/A	Yes
28	2	6	157,610,787	6	157,691,692	80,905	INV	3		N/A	Yes
28	2	8	39,433,540	8	39,645,931	212,391	INV	2		N/A	Yes
28	2	8	41,101,130	8	94,940,312	53,839,182	ITX	2		N/A	Yes
28	2	8	39,890,169	8	42,153,400	2,263,231	INV	3		N/A	Yes
28	2	8	39,903,411	8	42,119,649	2,216,238	INV	3		N/A	Yes
28	2	8	90,065,770	8	94,299,420	4,233,650	INV	3		N/A	Yes
28	2	8	40,087,838	8	119,896,781	79,808,943	INV	4		N/A	Yes
28	2	8	42,219,254	8	42,537,282	318,028	DEL	4		N/A	Yes
28	2	8	107,232,675	8	115,664,954	8,432,279	INV	4		N/A	Yes
28	2	8	42,684,185	8	47,378,680	4,694,495	DEL	5		N/A	Yes
28	2	8	41,352,306	8	119,468,390	78,116,084	DEL	7		N/A	Yes
28	2	8	108,063,856	8	119,962,758	11,898,902	INV	7		N/A	Yes
28	2	11	61,841,815	14	81,784,103	N/A	CTX	5		N/A	Yes
28	2	12	28,142,935	12	28,233,850	90,915	INV	10		N/A	No
28	2	12	15,728,835	12	27,920,974	12,192,139	INV	11		N/A	No
28	2	12	28,149,131	12	28,208,938	59,807	INV	11		N/A	No
28	2	12	4,347,749	12	4,388,753	41,004	DEL	12		N/A	No
28	2	12	15,766,773	12	28,086,973	12,320,200	ITX	12		N/A	No
28	2	12	4,355,223	12	27,836,020	23,480,797	INV	13		N/A	No

28	2	12	4,209,379	12	28,343,467	24,134,088	ITX	15		N/A	No
28	2	12	27,824,315	12	27,981,093	156,778	DEL	16		N/A	No
28	2	12	4,434,628	12	43,476,441	39,041,813	DEL	23		N/A	No
28	2	12	15,797,051	12	27,666,108	11,869,057	DEL	26		N/A	No
28	2	22	22,700,813	22	22,752,223	51,410	ITX	2		N/A	Yes
29	2	1	850,051	1	3,634,221	2,784,170	INV	3		N/A	Yes
29	2	2	201,406,407	13	59,640,949	N/A	CTX	2		N/A	No
29	2	2	218,567,435	2	220,713,459	2,146,024	INV	3		N/A	Yes
29	2	3	36,323,064	3	46,366,130	10,043,066	DEL	2		N/A	Yes
29	2	3	46,372,436	3	49,153,472	2,781,036	INV	4		N/A	Yes
29	2	3	36,872,783	3	38,170,567	1,297,784	ITX	7		N/A	Yes
29	2	13	26,523,960	13	85,856,260	59,332,300	DEL	2		N/A	No
29	2	13	33,090,332	13	59,042,060	25,951,728	INV	2		N/A	No
29	2	13	33,740,272	13	54,529,373	20,789,101	INV	2		N/A	No
29	2	13	36,751,832	13	96,892,395	60,140,563	INV	2		N/A	No
29	2	13	47,855,840	13	112,148,331	64,292,491	INV	3		N/A	No
29	2	13	47,865,586	13	49,013,788	1,148,202	INV	3		N/A	No
29	2	13	63,859,980	13	67,311,568	3,451,588	INV	3		N/A	No
29	2	13	69,202,753	13	85,543,411	16,340,658	DEL	3		N/A	No
29	2	13	77,956,326	13	81,861,275	3,904,949	ITX	3		N/A	No
29	2	13	78,877,239	13	113,321,390	34,444,151	INV	3		N/A	No
29	2	13	43,242,545	13	63,849,050	20,606,505	DEL	4		N/A	No
29	2	13	70,834,368	13	75,263,383	4,429,015	INV	4		N/A	No
29	2	13	75,290,525	13	97,029,030	21,738,505	INV	4		N/A	No
29	2	13	85,198,141	13	106,946,020	21,747,879	ITX	7		N/A	No
29	2	13	85,831,787	13	96,899,055	11,067,268	DEL	7		N/A	No
29	2	13	85,534,542	13	87,721,266	2,186,724	INV	8		N/A	No

[#]The reported amplitude may be a result of more than one rearrangements spanning the same area.

[*]ND Not done. Plasma DNA not analyzed due diffiulties in designing sensitive assays, due to low complexity regions or repeat regions or no more DNA to analyze

[$] Preserved in early, but not late metastasis

**N/A Not applicable. No tissue or plasma available