Cancer patient stratification by Next Generation Sequencing of Circulating Tumour DNA: Taking the next steps towards implementation in the clinic.

Dr Mark Openshaw

PhD

SUPERVISORS – PROFESSOR ANNE THOMAS & PROFESSOR JACQUI SHAW

NOVEMBER 2018

Abstract

Cancer patient stratification by Next Generation Sequencing of Circulating Tumour DNA: Taking the next steps towards implementation in the clinic.

Mark Openshaw

Recent advances in next generation sequencing (NGS) have allowed significant improvement in the sensitivity of detection of the tumour derived fraction of cell free DNA termed circulating tumour DNA (ctDNA). This thesis investigated the clinical application of ctDNA in serial plasma samples from breast cancer and gastroesophageal adenocarcinoma (GOA) patients undergoing palliative and curative treatment. Additionally, evaluation of tumour and ctDNA whole exome sequencing (WES) was completed on serial samples from one breast cancer patient.

Targeted NGS was used to detect tumour specific mutations. The same technology and commercial/designed droplet digital assays were used to detect ctDNA. WES was undertaken in Spain and the resultant sequencing data analysed in-house.

Ampliseq™ NGS detected ctDNA in 9 of 13 patients with metastatic breast cancer (69.2%). Detected mutations tracked disease progression and response. Oncomine™ NGS detected ctDNA in 5 out of 10 patients with pre and post neoadjuvant chemotherapy plasma samples. Two ctDNA positive patients relapsed, while none of the 5 ctDNA negative patients relapsed.

WES of ctDNA was successfully undertaken. Mutations detected from ctDNA WES overlapped with WES of the primary and relapsed tumours. Key tumour mutations could be identified in ctDNA including ongoing mutational changes caused by tumour evolution.

ctDNA was detected in a cohort of 44 patients with GOA. Nine of 13 (69.2%) metastatic patients with a detectable tumour mutation had detectable ctDNA. ctDNA levels tracked disease response and high levels of ctDNA at diagnosis of metastatic disease predicted short survival. Seven of 19 patients (36.8%) treated with curative intent that had a detectable tumour mutation had detectable ctDNA. Three patients had detectable ctDNA post surgery, two of whom relapsed during follow-up. Detection of ctDNA at any time prior to relapse predicted poor relapse free survival (HR 21.6, CI 2.5-185.5, p=0.005).

This thesis demonstrates the potential clinical utility of ctDNA in both breast cancer and GOA and shows an emerging role for the use of ctDNA WES.

Acknowledgments

I would like to thank Prof Anne Thomas and Jacqui Shaw for the opportunity to join them and undertake such a fascinating period of research. I enjoyed meeting every member of the lab but would most like to thank:

- Dr Barbara Ottolini for teaching and supporting me through my first year and a half, your scientific knowledge and insight was invaluable.
- Soon to be Dr Luke Martinson for his humour and support as a friend and a scientist as we both completed our PhDs, may the future ski trips be numerous and enjoyable.
- Drs David Guttery, Karen Page, Daniel Fernandez-Garcia and Lindsay Primrose for all their scientific knowledge and supply of cakes over the years.

To my sister and mum, thank you for all the support and holidays that kept me sane during my years at Leicester, and to my nephews Joe and Jamie and niece Annabelle for all the fun downtimes. To all my friends who were there when they were needed, I look forward to having more time to spend with you now.

Finally, to the people at 'Hope Against Cancer' in particular Barbara North and for Nigel Rose for funding the PhD, supporting all my talks, and pushing me to reach my exercise limits on the Hope to hope cycle ride.

Related Publications and Posters

Publications

1. OPENSHAW, M. R., PAGE, K., FERNANDEZ-GARCIA, D., et al. 2016. The role of ctDNA detection and the potential of the liquid biopsy for breast cancer monitoring. *Expert Rev Mol Diagn*.

2. OPENSHAW, M. R., RICHARDS, C. J., GUTTERY, D. S., et al. 2017. The genetics of gastroesophageal adenocarcinoma and the use of circulating cell free DNA for disease detection and monitoring. *Expert Rev Mol Diagn,* 17**,** 459-470.

3. OPENSHAW, M. R., HARVEY, R. A., SEIBRE, N. J., et al. 2015. Circulating Cell Free DNA in the Diagnosis of Trophoblastic Tumours. *EBioMedicine*, 29:4: 146-52.

Posters

1. OPENSHAW MR, OTTOLINI B, RICHARDS CJ GUTTERY D, SHAW JA, THOMAS A. Using the liquid biopsy to detect minimal residual disease in Gastroesophageal adenocarcinoma. ESMO 2018 congress. 19-23rd October 2018. Munich, Germany.

2. HASTINGS RM, OPENSHAW MR, VAZQUEZ M, FERNANDEZ-GARCIA D, GUTTERY D, PAGE K, TOGHILL B, THOMAS A, AHMED S, TOLEDO RA, SHAW JA. Whole-exome cfDNA profiling captures the mutational signatures of metastatic breast cancer for monitoring disease evolution. ESMO 2018 congress. 19-23rd October 2018. Munich, Germany.

3. OPENSHAW MR, OTTOLINI B, RICHARDS CJ GUTTERY D, SHAW JA, THOMAS A. Developing the liquid biopsy in gastroesophageal adenocarcinoma: Disease monitoring and detection of minimal residual disease. CNAPS 2017 20-22nd September 2017. Montpellier, France. and CRUK International Symposium on Oesophageal Cancer 27-28th April 2017. Cambridge, UK

Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Definition
AI	Aromatase inhibitor
ARID1A	AT-Rich Interaction Domain 1A
BAM	Binary alignment map (BAM)
bp	Base pair
BWA	Barrow-Wheeler aligner
cfDNA	Cell free DNA
CCNE1	Cyclin E1
CDH-1	Cadherin 1
CDKN2A	Cyclin Dependent Kinase Inhibitor 2A
CGI	Cancer Genome Interpreter
CI	Confidence Interval
CR	Complete Response
COSMIC	Catalogue of Somatic Mutations in Cancer
CSF	Cerebrospinal fluid
Ct	Cycle threshold
CT	Computed tomography
CTCs	Circulating tumour cells
ctDNA	Circulating tumour DNA
ddPCR	Droplet digital PCR
EGFR	Epidermal Growth Factor Receptor
ER	Oestrogen receptor
ERBB2	Erb-B2 Receptor Tyrosine Kinase 2
ESR1	Oestrogen receptor 1
FF	Fresh Frozen
FFPE	Formalin fixed paraffin embedded
FISH	Fluorescent in situ hybridisation
FRP	Forward Reverse Probe
GATK	Genome Analysis Tool kit
GOA	Gastroesophageal adenocarcinoma
GOJ	Gastroesophageal junction
HER2	Human epidermal growth factor receptor 2
ISP	Ion sphere™ particles
IGV	Integrative Genomics Viewer
JAK2	Januse kinase-2
KRAS	K-ras oncogene
MRI	Magnetic resonance imaging
MS	Mutational signature
MSI	Microsatellite instability
MYC	MYC Proto-Oncogene
ng	Nanogram
NGS	Next generation sequencing

OGD	Oesophago-gastroduodenoscopy
OS	Overall survival
PCR	Polymerase chain reaction
PD	Progressive disease
PD-L1	Programmed death ligand -1
PFS	Progression free survival
PGM™	Personalised genome machine
PIK3CA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha
PR	Partial Response
PTEN	Phosphatase and Tensin Homolog
qPCR	Quantitative PCR
ROC	Receiver Operating Characteristic
SAM	Sequence alignment map (file)
SCNA	Somatic copy number variant
SD	Stable Disease
SMAD4	SMAD Family Member 4
SNV	Single-nucleotide variant
SOX2	Sex determining region Y-Box 2
TP53	Tumour Protein P53
TP53BP2	*TP53* binding protein 2
µg	microgram
VAF	Variant allele fraction
WES	Whole exome sequencing
WGS	Whole genome sequencing

1 Introduction: Circulating Tumour DNA

1.1 Background

Circulating cell free DNA (total cfDNA) in the plasma was first observed in 1948 even before the structure of DNA was discovered (Mandel and Metais, 1948). As early as 1977 Leon et al. (1977) showed that patients with cancer had higher levels of total cfDNA than unaffected people and that there was a correlation between response to treatment and reduction in levels of total cfDNA . Sequencing of the *KRAS* gene in the plasma of patients with pancreatic cancer showed the first evidence for detection of mutated oncogenes that could be matched with the primary tumour (Sorenson et al., 1994). This showed that DNA derived from tumour cells, termed circulating tumour DNA (ctDNA), could be differentiated from total cfDNA derived from apoptosis of normal body cells and provide genetic information about the tumour. The ability of ctDNA to derive information about a tumour that was previously only available through a needle biopsy, led to the term "liquid biopsy" to describe ctDNA. Since blood is routinely obtainable and ctDNA is readily detectable notably in advanced disease (Bettegowda et al., 2014), investigation of the utility of ctDNA to act as a liquid biopsy is an important area of research. However, whether ctDNA analysis can act as a true surrogate for tissue biopsy depends on the extent to which it is present and detectable in patients with cancer and the extent to which it reflects genetic changes in the tumour over time and in response to interventions.

1.1.1 Origins of cell free DNA

Previous studies have shown that total cfDNA is generated predominantly by apoptosis and necrosis (Jahr et al., 2001) although active secretion (Stroun et al., 2001, Anker et al., 1975) may also occur, as total cfDNA can be detected in the supernatant of pre-implanted human embryos (Assou et al., 2014). ctDNA generated by necrosis is generally less fragmented

(fragments >10,000bp) than that generated via apoptosis, which generates fragments of 180bp or less (Jahr et al., 2001). This difference is hypothesised to occur due to the higher DNAase activity during controlled apoptotic cell death (Nagata, 2000). It should be noted that the majority of ctDNA is not likely to be derived from circulating tumour cells (CTCs), as CTCs are relatively rare in the circulation and ctDNA levels are often 100 to 1000 times the concentration that could be generated from apoptosis of CTCs alone (Thierry et al., 2016).

Overall the proportion of ctDNA in total cfDNA varies widely, ranging from <0.01% to sometimes in excess of 90% (Jahr et al., 2001) but these cases are unusual with a range of 0-65% more common in metastatic patients (Jung et al., 2010) and less in patients with localised tumours. Different cancers also appear to produce total cfDNA by different routes as can be seen by the pattern of DNA fragmentation in the serum (Jahr et al., 2001). This may contribute to differing degrees of ctDNA detection between cancer types as longer DNA fragments can more easily be amplified by PCR. Overall the ability to detect ctDNA depends on the ability to detect genetic alterations of the specific cancer within the pool of total cfDNA. As the ctDNA may make up <0.1% of the total cfDNA, highly sensitive methods for detection of ctDNA are required.

1.1.2 Tumour Heterogeneity and ctDNA

It is widely accepted that the tumours of patients with cancer show a significant degree of intratumoural heterogeneity (Gerlinger et al., 2012, de Bruin et al., 2014, Allott et al., 2016, Hanahan and Weinberg, 2011). It is believed that cancers acquire resistance to drug therapy through the process of clonal selection and evolution of resistant cells (Aparicio and Caldas, 2013) dependent upon this inherent heterogeneity. ctDNA has been shown to be present in the majority of patients with advanced and metastatic cancers (Bettegowda et al., 2014) and in some patients with localised cancers (Chen et al., 1999). DNA sequencing studies have suggested that ctDNA may be more representative of total tumour

heterogeneity in a patient (Chan et al., 2013) than single biopsies, that can only represent a single tumour region. This is because all tumours are believed to contribute to the ctDNA pool (Thierry et al., 2016). ctDNA analysis may therefore allow assessment of tumour heterogeneity at a greater frequency and with reduced morbidity compared to tissue biopsies. Advantages and disadvantages of the two types of analysis are shown in Figure 1.

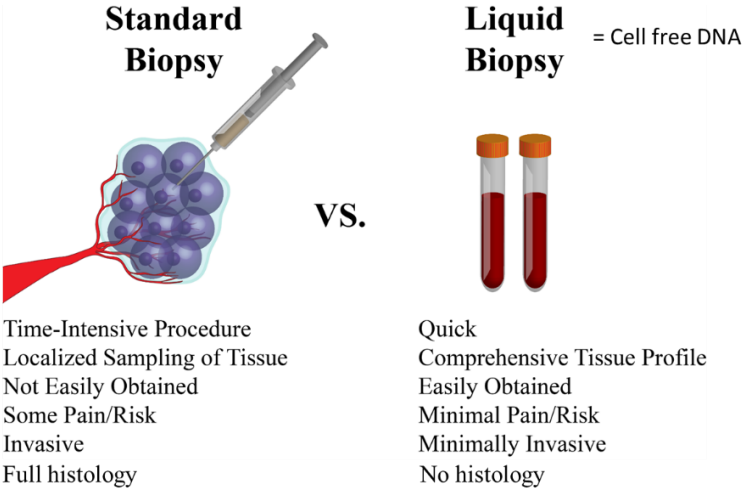


Figure 1 – **Advantages and disadvantages of standard and liquid biopsies.**

1.1.3 Methods of detection of circulating tumour DNA

1.1.3.1 Genome variability

In cancer, somatic genetic alterations are acquired in the DNA of tumour cells in response to genotoxic DNA damage and replicative errors (Hanahan and Weinberg, 2011). The genetic alterations can include single-nucleotide variants (SNVs), short base substitutions, chromosomal rearrangements and translocations, and somatic copy number variants (SCNAs) (Wang et al., 2004, Chen et al., 1999, Nik-Zainal et al., 2012, Iafrate et al., 2004). SCNAs refer to gain and loss of segments of DNA resulting in deletion of some chromosomal regions and amplification of others (Griffiths et al., 1999). Since these somatic genetic alterations are not present in normal body cells, detection of these tumour specific changes allow ctDNA to be distinguished from total cfDNA. In order to detect these genetic changes a variety of methods have been developed, each of which has different advantages and disadvantages.

1.1.3.2 Next generation Sequencing

The development of next generation sequencing (NGS) has allowed a step change in the analysis of tumour and cell-free DNA. NGS is used to describe the simultaneous (parallel) sequencing of many millions of DNA sequences 50-500 base pairs (bp) in length. In comparison to Sanger sequencing that uses chain termination followed by capillary based fragment separation, NGS enables determination of DNA sequence by DNA synthesis where nucleotides are added prospectively using a variety of techniques including bead and flow cell based methods (Bahassi el and Stambrook, 2014).

Hotspot mutation panels allow sequencing of a limited number of specific genes and are frequently referred to as targeted NGS. Prior to sequencing the DNA pool is enriched for specific genes or gene regions, called amplicons, by a number of methods including PCR or oligonucleotide-based in-solution capture methods (Parla et al., 2011). These targeted sequencing panels provide high sensitivity detection of variants and are capable of picking up many of the low frequency alterations, <5% variant allele fraction (VAF), typical of ctDNA (Okamoto et al., 2014, Kim et al., 2015). As a result, commercial hotspot panels are readily available and are frequently tailored to the genetic alterations of specific cancers.

The limit of sensitivity of amplicon based targeted (hotspot) NGS panels has improved over the last 3-4 years. The introduction of panel assays with high average read depth (defined as the number of times each base in a gene is read) and that use unique molecular barcodes (Kivioja et al., 2011) to enable error suppression, has reduced the limit of sensitivity from a VAF of ≥1% to ~0.1% (Guttery et al., 2015, Vollbrecht et al., 2018, Gale et al., 2018, Lanman et al., 2015, Newman et al., 2014). However, the true limit of sensitivity for each panel is dependent on the mutation(s) covered, panel complexity and amount of starting DNA template.

Whole exome sequencing (WES) is an NGS approach that involves sequencing the entire coding region of the genome (30,000+ genes). Although this makes up only 1% of the

genome, mutations in coding regions are understandably frequently linked to disease. Recently costs of WES have greatly reduced and therefore it is often used to analysis tumour genetics (Ross and Cronin, 2011). However, due to the increased breadth of gene coverage in WES there is reduced read depth which raises the limit of sensitivity for mutation detection. Other limitations of WES include the cost of processing the large amount of information generated and resultant slow analysis times. For this reason, WES has been infrequently used in ctDNA analysis, and its use has only been recently reported (Butler et al., 2015).

1.1.3.3 Droplet Digital PCR

Droplet digital ddPCR is an emulsion-based technology technique whereby thousands of separate PCR reactions are carried out simultaneously within tiny droplets of PCR mixture suspended in oil. The DNA variations are detected using fluorescently tagged Taqman® probes (see methods 2.12) which are mutation specific and cause each droplet containing mutant DNA to fluoresce differently to droplets containing wildtype DNA. These droplets are then counted by flow cytometry (Hindson et al., 2011). This has greater sensitivity of mutation detection than NGS, down to a 0.01% VAF with 10ng template DNA. However only one or a few mutations can be detected in a single assay, and as the assays are mutation specific the mutation must already be known (e.g. *BRAF* V600E in malignant melanoma) to use this method (Hudecova, 2015). In contrast NGS will detect any mutation within the amplicons contained with a panel.

1.1.3.4 Copy number Variation

Whilst NGS is useful for detecting somatic mutations and in particular SNVs, it is less frequently used for SCNA analysis (Hehir-Kwa et al., 2015). Traditional methods of tumour DNA analysis include microarrays, but more recently other methods have been utilised, including the Nanostring nCounter®, which uses a highly multiplexed fluorescent probe

based system (Figure 2) (Veldman-Jones et al., 2015). However, this requires high DNA input amounts (300ng), which make total cfDNA analysis challenging when the yield is low.

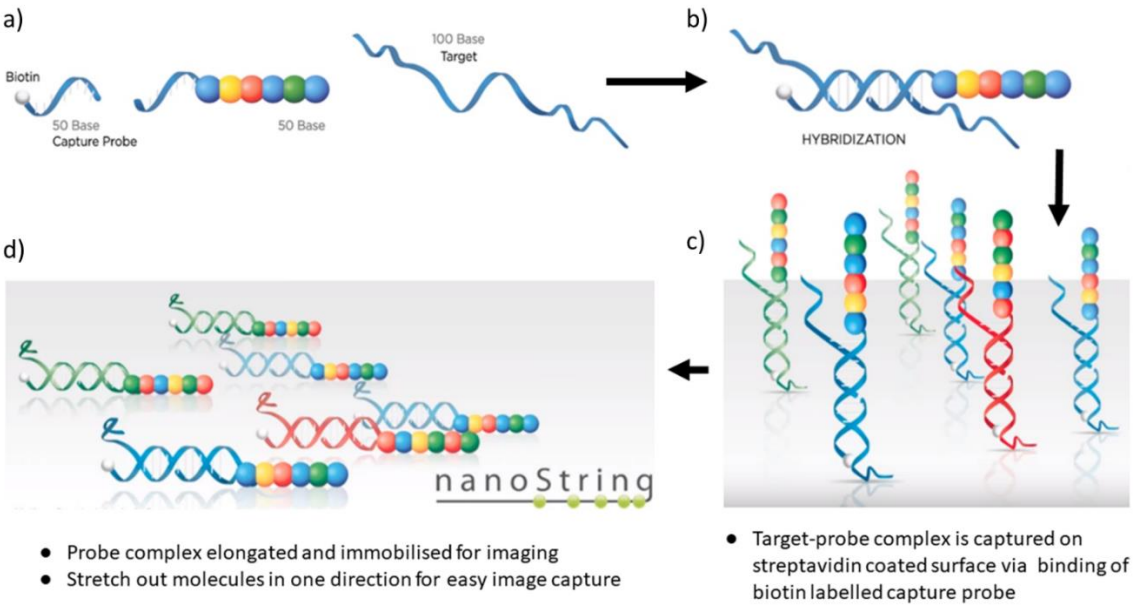


Figure 2 **– Nanostring nCounter technology overview. a)** 50bp capture probe and 50bp reporter probe b) binding of capture and reporter probe to 100bp target sequence c) target probe complex is captured on streptavidin surface d) imaging of bound probes.

Quantitative PCR (qPCR) can detect SCNA using much smaller amounts of template DNA (<10ng), allowing total cfDNA analysis. This method involves relative quantitation (RQ), or ratio, of the PCR product of the gene of interest to a standardised non-duplicated reference gene, or preferably the mean of a number of reference genes, to detect gene amplification (Page et al., 2011, Shoda et al., 2014). Primers for the gene(s) of interest are designed along with a Taqman® fluorescent probe which allow detection of PCR amplification in real time. ddPCR can also be used to quantify amplification (Shoda et al., 2016) as it relies on the same PCR (Taqman®) technology to generate gene ratios. However, as with SNVs, the amplification that is analysed must already be known as both qPCR and ddPCR SCNA assays are gene specific. Unfortunately, it is very difficult to detect gene deletion via this method.

1.1.4 Dynamic nature of ctDNA and disease monitoring

The less invasive nature of plasma sampling (compared to needle biopsy) allows frequent blood sample collection for monitoring of cancers. Studies have shown that levels of ctDNA reduce rapidly following surgical removal of tumours (Diehl et al., 2008, Khier and Lohan, 2018) showing that ctDNA levels are dynamic, and respond quickly to changes in tumour load, indicating a significant role for ctDNA in disease monitoring. This role has been confirmed following analysis of serial blood samples from patients with metastatic colorectal, breast, lung and other cancers (Diehl et al., 2008, Dawson et al., 2013, Hamakawa et al., 2014, Ni et al., 2017). In these studies, rising levels of ctDNA have been shown to predict disease progression, whilst falling levels track disease response to therapy. In further studies, changes in the mutations present in ctDNA give clues to the mechanism of resistance to anticancer therapies (Thierry et al., 2014, Li et al., 2018a, Guttery et al., 2015). The dynamic nature of ctDNA is partially due to the short half-life of total cfDNA in the plasma, calculated at 15-120 minutes (Khier and Lohan, 2018, Lo et al., 1999) but also due to the underlying mechanism of production, which is linked to the level of cancer cell turnover and tumour burden (Thierry et al., 2016). Furthermore, in early stage disease, the appearance of ctDNA during blood monitoring following surgery has been shown to predict relapse and detect changes in the genetic make-up of the tumours (Garcia-Murillas et al., 2015, Abbosh et al., 2017). These studies show that ctDNA can measure temporal as well as spatial tumoural heterogeneity. They also indicate that ctDNA has clinical utility in monitoring disease in both adjuvant and metastatic patients and, since ctDNA is detectable in early stage disease, there is also a developing role for ctDNA in the screening of high risk populations (Cohen et al., 2018).

1.1.5 Overview

With the development of new technologies, techniques and understanding of the relationship between genetic alterations in tumours and ctDNA, interest has rapidly grown into the utility of ctDNA detection. There are multitude of genetic changes that occur in tumours that can also be detected in ctDNA. This means ctDNA may have a role in disease

monitoring, detection of resistance to anti-cancer therapy, prediction of relapse and potentially aiding diagnosis and screening. ctDNA analysis may therefore revolutionise the monitoring and management of wide range of cancer patients (Shaw and Stebbing, 2014). However, the use of ctDNA in the clinic using commercial tests is still at its early stages and can attract significant controversy (Torga and Pienta, 2018, Thoma, 2018), caused mainly by poor understanding of the technology and interpretation of results. In order to become widely used in the clinic, the application and utility of ctDNA must be understood by a wide range of clinicians and the benefits of its use demonstrated in large scale clinical trials.

1.2 An introduction to breast cancer and ctDNA

1.2.1 Background

Breast cancer is the most common cancer in women worldwide, with 1.7 million new cases diagnosed in 2012 alone (Ferlay et al., 2013). Significant advances have been made in the field of neoadjuvant and adjuvant therapy such that breast cancer mortality has been declining in the majority of developed countries (Bosetti et al., 2012, Jatoi et al., 2007). Selecting patients that benefit most from adjuvant therapy is classically based on clinico-pathological features such as stage, tumour size, lymph node involvement, grade and receptor status (Weigelt et al., 2005). Where these features provide insufficient certainty on treatment decisions, the use of multi-gene panel analysis of primary tumours such as Oncotype Dx can further stratify decision making, highlighting the importance of understanding tumour genetics (Sparano et al., 2018).

OncoType Dx is a 21 gene panel that provides prognostic information in node negative oestrogen receptor (ER) positive, human epidermal growth factor receptor 2 (HER2) negative cancers (SanGiovanni et al., 2015). OncoType Dx and other similar tests rely on information from tissue at a single time point and, apart from at surgery for neoadjuvant therapy, no further pathological assessment of tumours is undertaken during curative therapy, leaving a gap in our understanding of tumour biology. Following surgery with curative intent it is not currently possible to discriminate those patients who have minimal residual disease and are at risk of relapse from those who are not. In addition, relapses may occur many years from initial treatment causing significant patient concern (Fisher et al., 2002). A biomarker that could detect minimal residual disease and monitor genetic changes in response to therapy would therefore significantly improve the management options for patients with breast cancer.

1.2.2 Monitoring tools in metastatic breast cancer

The management of metastatic breast cancer requires the use of a variety of scans, tumour analyses and blood tests to monitor disease response and progression and guide treatment

selection. Traditional monitoring tools include computed topography (CT) scanning and the serum biomarker cancer antigen 15-3 (CA 15-3). CTCs have also been investigated as a way of monitoring breast cancer (Hayashi and Yamauchi, 2012). Dawson et al. (2013) compared these three modes of monitoring disease with total cfDNA analysis. They used targeted and whole genome sequencing (WGS) to identify unique genetic alterations in the tumours of 30 different women with metastatic breast cancer. They detected the same somatic mutations in total cfDNA, confirming the presence of ctDNA. This ctDNA was detected in a greater proportion of patients and had a greater correlation with tumour burden than either CA15-3 or circulating tumour cells. Although an expensive and time consuming way of monitoring disease due to the initial need to identify the unique genetic alterations in each cancer patient, it confirmed the potential for total cfDNA analysis to be an informative and sensitive biomarker in metastatic breast cancer.

1.2.3 Tumour heterogeneity in breast cancer

Most breast cancer patients initially present with localised breast cancer and go on to receive treatment with curative intent. It is now recommended that all patients who present with relapsed metastatic disease have a repeat biopsy where there is a lesion amenable to biopsy (Cardoso et al., 2014) because the ER, progesterone receptor and HER2 status of the metastatic tumour may be discordant with that of the primary tumour (Curtit et al., 2013). This degree of temporal heterogeneity is not restricted to receptor status and may be seen in many different genes (Shah et al., 2009). Breast cancers also display significant spatial heterogeneity (Nik-Zainal et al., 2012, Shah et al., 2012) as seen in other cancers (Gerlinger et al., 2012). In some patients, biopsy of disease may not be feasible due to disease location and even where it is feasible, spatial heterogeneity determines that a single biopsy is unlikely to represent the full range of genetic alterations in a patient's tumours. De Mattos-Arruda et al. (2014) investigated the ability of ctDNA to interrogate tumour genetics by undertaking a form of targeted NGS sequencing covering 300 known cancer genes. They analysed the primary breast tumour, liver metastatic deposit and total cfDNA of a patient who presented with metastatic breast cancer. Not all mutations

identified in the metastasis were identified in the primary tumour supporting the notion of intratumoural heterogeneity. More interestingly total cfDNA analysis captured all mutations present in the primary and metastatic tumours, demonstrating the ability of ctDNA to reflect patient wide intratumoural heterogeneity.

Rothe et al. (2014) investigated the potential of ctDNA to act as an alternative to metastatic biopsies for mutational analysis in 17 patients with metastatic breast cancer. They used the Ion AmpliSeq™ cancer hotspot panel covering mutations in 50 cancer genes. They found a high degree of concordance (76%, Confidence interval (CI) 50-93%) between tumour and total cfDNA collected at the same time point. In 2 cases where a mutation was identified in the tumour but not the plasma they went on to detect the alterations in the plasma using an alternative NGS technology. In the two cases where mutations were present in total cfDNA but not the tumour DNA the mutant allele frequencies were sufficiently high to indicate (>10% VAF) that it was likely these mutations had been missed by the metastatic biopsy due to intratumoural heterogeneity.

These studies highlight the ability of ctDNA to act as an informative and sensitive biomarker and the potential for ctDNA to function as a "liquid biopsy", which reflects body-wide intratumoural heterogeneity more accurately than a single tissue biopsy. ctDNA may therefore provide mutational analysis of important cancer genes that can be utilised as an alternative to metastatic biopsies in molecular screening programs.

1.2.4 Guiding current therapies

1.2.4.1 Mutation (SNV) tracking and endocrine therapy

A study by Guttery et al. (2015) used total cfDNA analysis to provide clinically useful information. Using targeted NGS, a gene panel was created for mutations in 5 known cancer genes covering 23 amplicons. This included phosphatidylinositol-4,5-Bisphosphate 3-Kinase, Catalytic Subunit Alpha (*PIK3CA)*, tumour protein 53 (*TP53*) and fibroblast growth factor receptors 1 and 2 (*FGFR1 & FGFR2*) which are known to be mutated in breast cancer

but for which no effective therapeutic intervention is routinely available (Nik-Zainal et al., 2016). They also analysed the oestrogen receptor 1 (*ESR1*) gene. Recent work has shown that oestrogen receptor α activating mutations are acquired in approximately 20% of patients on endocrine therapy as a result of *ESR1* mutations (Jeselsohn et al., 2014). These mutations cause ligand independent oestrogen receptor transcriptional activity, which is unresponsive to endocrine manipulation, resulting in resistance to currently available anti-oestrogen therapies such as tamoxifen and aromatase inhibitors (AIs) (Jeselsohn et al., 2014). Specifically the two most common *ESR1* mutations p.D538G and p.Y537S have been linked with activation of the Insulin-like growth factor (IGF) pathway resulting in increased growth stimulation in cell line assays (Li et al., 2018b). As with Rothe et al. (2014), Guttery et al. (2015) found mutations in *PIK3CA* and *TP53* in the tumour and total cfDNA. In addition, *ESR1* mutations were found in the total cfDNA of 15 patients and using both NGS and ddPCR. These patients were therefore predicted to be resistant to the commonly available anti-oestrogen therapies despite being ER positive. It was proposed that this technique could detect *de novo* mutations in *ESR1* during anti-oestrogen therapy and therefore allow longitudinal monitoring for therapy resistance, a clinically useful goal in breast cancer. The detection of *ESR*1 mutations in total cfDNA has since been confirmed in a larger cohort of 80 patients (Beije et al., 2018) but has yet to enter routine clinical use.

More recently O'Leary et al. (2018) showed that reduction of *PIK3CA* mutation levels in total cfDNA, in the PALOMA-3 study, provided a robust biomarker for prediction of response to the CDK4/6 inhibitor palbociclib plus fulvestrant therapy in metastatic breast cancer. A reduction in ctDNA levels (as measured by *PIK3CA* copies/ml) from day 1 to 15 was a strong positive predictor of disease response.

In summary, the identification of resistance to anti-oestrogen therapy by detection of *ESR1* mutation in total cfDNA has been widely shown. Prospective identification of *ESR1* mutations is predicted to allow switching to other treatments before radiographic disease progression. In addition, monitoring ctDNA levels using truncal mutations such as *PIK3CA* (and by extrapolation of *TP53*) may also predict response to endocrine therapies.

1.2.4.2 Human Epidermal Growth Factor Receptor 2 (HER2)

The Erb-B2 Receptor Tyrosine Kinase 2 (*ERBB2)* gene that encodes the HER2 protein plays a crucial role in regulation of cell growth and is overexpressed in 20-30% of breast cancers (Hudis, 2007). The current routine method for detecting HER2 status is the detection of expression via immunohistochemistry, resulting in a score of 0-3+, plus the detection of amplification via fluorescent in situ hybridisation (FISH) for patients that are HER2 2+. Patients scored 0-1 and 2+ FISH negative, are considered non-responsive to HER2 targeted therapy, whilst 3+ and 2+, FISH positive are considered responsive (Rakha et al., 2015). This test requires sufficient tumour biopsy tissue to be available. Detection of *ERBB2* amplification alone can provide an alternative way of assessing responsiveness and suitability for trastuzumab therapy (Vogel et al., 2001) but is not widely used. Page et al. (2011) have shown that qPCR can be applied to tumour DNA and total cfDNA from patients with breast cancer using a standard reference gene to detect *ERBB2* amplification. This assay could detect *ERBB2* amplification in 95.2% of HER2 overexpressing tumours. Amplification could also be detected in total cfDNA in 5/18 metastatic patients and 6/39 primary patients. These primary patients included postsurgical patients, many of whom may have been cured by therapy and therefore would not have been expected to have detectable *ERBB2* amplification. *ERBB2* gene amplification could also be seen in two HER2 negative patients post primary treatment suggesting that ctDNA analysis was detecting the development of *ERBB2* amplification. This is in keeping with previous studies showing that patients with HER2 negative tumours can develop into HER2 positive tumours over time (Meng et al., 2004). The detection of *ERBB2* amplification in the total cfDNA of patients who had completed curative treatment and had no evidence of disease suggests that this technique was detecting minimal residual disease. Whether these patients went onto relapse is not known but is an important factor to determine in future studies. The study shows the utility of total cfDNA analysis in determining *ERBB2* status in a subset of HER2 positive patients and may therefore allow selection of HER2-targeted therapies in the future.

1.2.5 Predicting relapse / minimal residual disease

Evidence for the utility of total cfDNA to detect minimal residual disease has been published by Garcia-Murillas et al. (2015). This was a study of 55 neoadjuvant breast cancer patients, who had total cfDNA analysis at diagnosis, post surgery and then at 6 monthly intervals. They first detected somatic mutations (primarily in *PIK3CA* and *TP53*) in the primary tumour by targeted NGS of 14 known breast cancer driver genes. They then used personalised ddPCR assays to track each identified mutation. They showed that the presence of ctDNA in the post-surgical sample (defined by detectable mutations via ddPCR) was strongly correlated with a risk of relapse. In 6 of 7 patients that relapsed during the study, ctDNA predicted relapse giving a hazard ratio of 25.1 (95% CI 4.08-130.5) for ctDNA positivity. The use of multiple follow-up samples further increased the ability to predict relapse. An increase in ctDNA levels predicted relapse a median of 8 months prior to metastatic relapse. This is a similar correlation to that seen in *EGFR* and *KRAS* mutations in colorectal cancer (Diaz et al., 2012) and in previous studies of total cfDNA and breast cancer (Silva et al., 2002). Of greater interest Garcia-Murillas et al. (2015) used high depth total cfDNA sequencing to characterise the genetics of the minimal residual disease, and compared it to both the primary tumour and available metastatic biopsies. The sequenced ctDNA showed more similarities with the metastatic biopsies than the original primary and in one case identified the loss of an *ESR1* mutation. This loss of an *ESR1* mutation predicted ER loss in the metastasis which was subsequently confirmed. This shows that ctDNA can predict those at high risk of disease relapse and in practice, determine resistance mutations in minimal residual disease many months before clinical progression. This shows that ctDNA is a valuable liquid biopsy in the adjuvant setting before metastatic biopsy is available.

Olsson et al. (2015) investigated evidence for minimal residual disease detection in ctDNA analysis of 20 post surgical patients. They looked for chromosomal rearrangements in the primary tumour via WGS. They then detected these rearrangements in the plasma using ddPCR to increase sensitivity of detection. They designed ddPCR assays for 4-6 chromosomal rearrangements per patient. Strikingly they detected ctDNA in 13 out of 14

patients that went on to relapse (93%) and in none of the 6 patients (0%) who had long term disease free survival. Detection of ctDNA preceded clinical detection of metastasis in 86% of patients, on average 11 months prior to relapse (range 0-37 months). This further supports the idea that total cfDNA analysis can detect genetic changes in minimal residual disease in advance of the development of overt metastatic disease.

It is known that breast cancer relapses can occur over 10 years after the primary cancer was initially detected and treated (Fisher et al., 2002). A dynamic accumulation of new mutations and loss of other genetic changes has been shown to occur over time in cancers (Gerlinger et al., 2012, Diaz et al., 2012). It is speculated that a break in cancer dormancy may occur after the collection of sufficient new mutations that allow renewed cell division, tumour growth and invasion to occur (Wan et al., 2013). Therefore, a sudden increase in the amount of detectable ctDNA may herald a break in dormancy of the disease and present a window of opportunity for treatment of minimal residual disease prior to the development of overt metastatic disease. If the detection of minimal residual disease can be combined with *HER2* and *ESR1* analysis along with future targeted therapies linked to specific somatic genetic alterations, the potential for successful treatment of these 'ctDNA detected pre-clinical relapses' exists. This has a theoretical chance of eliminating the minimal residual disease and suggests an exciting future for total cfDNA.

1.2.6 Gene specific total cfDNA analysis

The investigations discussed thus far used WGS or targeted sequencing to first define specific mutations in a patient, which are then further interrogated in the total cfDNA via NGS, ddPCR or qPCR. One method to avoid the initial stage of tumour analysis is to analyse common mutations in frequently mutated genes. Oshiro et al. (2015) investigated 3 common mutations in *PIK3CA*, which is mutated in 20-40% of breast cancers, (Stephens et al., 2012) via ddPCR. The lack of a screening stage and proven sensitivity for PIK3CA mutations of 0.01% means that this technique could be used for high patient throughput analysis. In the study 33% (110/330) of patients had *PIK3CA* mutations detectable in their

primary tumours via real time qPCR and 22.7% (25/110) had detectable *PIK3CA* mutations in total cfDNA. Although recurrence free survival (92% vs 68%, p=0.0029) and overall survival (96%vs 84% p=0.0283) were significantly shorter in ctDNA positive patients there was no significant increase in the relapse rate. However, when the ctDNA was dichotomised into $ctDNA^{high}$ and $ctDNA^{low}$ levels, the $ctDNA^{high}$ levels remained significant as a poor prognostic factor. This study used serum samples, whereas plasma samples are typically believed to be better for total cfDNA analysis. Despite this weakness this study highlights the ability of total cfDNA analysis to interrogate specific genetic mutations of a primary tumour without the need for a screening stage.

1.2.7 Breast Cancer and ctDNA Summary

ctDNA has been shown to be detectable in a wide range of human cancers (Openshaw et al., 2016, Diaz et al., 2012, Bettegowda et al., 2014, Manier et al., 2018, Cohen et al., 2018). In the case of metastatic breast cancer, ctDNA analysis has been proven to track tumour burden more accurately than currently available biomarkers and provide a method for tracking genetic changes over time (temporal heterogeneity) without the need for repeated biopsies. This promises to improve our understanding of the dynamic nature of tumour genomes. In addition, it may provide a method for detecting disease resistance and serve as a basis for changing anti-cancer therapies, including currently available therapies, before the advent of clinical progression. This is important because clinical progression detectable on CT scans represents the addition of many billions of cancer cells (Thierry et al., 2016), which will inevitably result in the accumulation of further genetic abnormalities and potential therapeutic resistance mutations. It is also important because it is the increase in tumour size that will eventually result in patient death and therefore use of the ctDNA to direct change in anti-cancer therapy before tumour growth, may be able to extend life in the metastatic setting. In the adjuvant setting total cfDNA analysis has been shown to detect minimal residual disease following surgery and to predict the development of metastatic relapse. Interrogation of ctDNA mutations prior to relapse may be useful to select appropriate therapies to delay clinical relapse and potentially destroy the residual

malignant cells thereby increasing the cure rate for these patients, which is, after all a highly desirable aim for a most promising developing technology.

1.3 An introduction to gastroesophageal adenocarcinoma and ctDNA

1.3.1 Background

Incidence rates for cancer of the stomach and oesophagus are high with 1.4 million cases diagnosed in 2012 (Ferlay et al., 2013). The majority of these cancers are gastroesophageal adenocarcinoma (GOA), which includes oesophageal adenocarcinoma (OAC), gastroesophageal junction (GOJ) adenocarcioma and stomach cancer. GOA is most common in eastern Asia with the highest prevalence rates in Japan and South Korea. However, rates of OAC and GOJ adenocarcinomas in the Western world have shown a recent significant increase (Thrift, 2016, Hartgrink et al., 2009). Overall survival rates are poorer than most other cancers with 1.1 million deaths worldwide in 2012 (Ferlay et al., 2013) and a five year survival of just 20-25% (Hartgrink et al., 2009, Thrift, 2016). There are limited treatment options for GOA with combination chemotherapy regimens such as EOX (Epirubicin, Oxaliplatin, Capecitabine) showing the best first-line activity in advanced disease. Despite this improvement in treatment median overall survival remains less than 12 months for advanced disease (Cunningham et al., 2008). There is therefore great clinical need to improve our understanding of these cancers to develop improved management and treatment strategies.

1.3.2 Is gastroesophageal adenocarcinoma one disease entity?

1.3.2.1 Traditional classification of gastroesophageal adenocarcinoma

GOA can be defined based upon anatomical origin, including tumours of the lower oesophagus, gastroesophageal junction (GOJ) and stomach. The TMN 8th edition classification system defines only two groups of cancers from this location; oesophageal and gastric tumours. GOJ tumours fall mainly within the oesophageal tumour grouping (Siewart types I and II) (Brierley et al., 2016.). However the falling frequency of gastric adenocarcinomas along with a rise in the incidence of GOJ adenocarcinomas (Devesa and Fraumeni, 1999) has led to uncertainty about the classification of gastro-oesophageal tumours, particularly those originating at or close to the GOJ (Hayakawa et al., 2016). As

our understanding of the biology of adenocarcinoma of the stomach and oesophagus improves, the divisions between these two groups have become less well defined and are therefore discussed in further detail.

1.3.2.2 Oesophageal adenocarcinoma and Barrett's oesophagus

Since the 1980's the incidence of lower OAC has increased in the United States and northern Europe. OAC is now more common than oesophageal squamous cell carcinoma in northern Europe with the highest incidence in Scotland, UK (Castro et al., 2014). These changes in frequency are thought to be caused by changes in diet, smoking, obesity and gastro-oesophageal reflux disease (GORD) rates (Lagergren and Lagergren, 2013, Thrift, 2016). Men are also disproportionately more affected by this disease than women (4.99 per 100,000 men vs 1.09 women)(Castro et al., 2014).

Barrett's oesophagus is the only well-defined precursor of OAC and GOJ adenocarcinoma. Barrett's oesophagus describes a condition in which squamous epithelium of the oesophagus is replaced by metaplastic columnar epithelium. The most important risk factor for Barrett's oesophagus is GORD causing oesophagitis, which in the long term leads to the development of Barrett's oesophagus (Lagergren and Lagergren, 2013). The metaplastic columnar epithelium of Barrett's oesophagus described as intestinal metaplasia, is predisposed to develop areas of adenocarcinoma occurring at a rate of approximately 0.33% per year (Desai et al., 2012) .

1.3.2.3 Gastric Adenocarcinoma and H.Pylori

Gastric adenocarcinoma is the second most common cause of cancer death worldwide (Ferlay et al., 2013). In Japan and South Korea surveillance programs have been developed due to the high prevalence of disease (Sugano, 2015) and have been effective at detecting early stage disease, but in the rest of the world including Europe, approximately two thirds of patients will present with advanced disease (Wagner et al., 2010). In addition, a

significant proportion of those treated with curative intent will relapse (Wagner et al., 2010). For patients with incurable disease combination chemotherapy improves survival by approximately 6 months and trastuzumab has been shown to be beneficial in patients overexpressing HER2 (Wagner et al., 2010). Trastuzumab represents the only currently available molecularly targeted therapy in gastric cancer, although other targets such as FGFR are under investigation (Hong et al., 2013). The aetiology and clinicopathological features of the disease also differ between Asian and Western patients (Kim et al., 2013) indicating that the disease may have genetic differences between the two locations.

Helicobacter Pylori (H. Pylori) is the strongest known risk factor for the development of gastric adenocarcinoma, with over 75% of gastric cancers being linked to this infection (Herrera and Parsonnet, 2009). It causes a persistent gastritis in infected people and can cause both diffuse type gastric-adenocarcinoma or the more prevalent intestinal type adenocarcinoma. Intestinal type adenocarcinoma is known to progress through various premalignant stages including atrophic gastritis and dysplasia to intestinal metaplasia and eventual adenocarcinoma, however only a small proportion of people infected with *H.pylori* will go on to develop cancer (Polk and Peek, 2010). EBV infection is a second potential risk factor that is currently being investigated (Akiba et al., 2008) whilst smoking and obesity are also described risk factors linked with OAC (Hartgrink et al., 2009) (Thrift, 2016).

1.3.2.4 Tumours of the gastro-oesophageal junction

Tumours of the GOJ have previously been classified depending upon their point of origin using the Siewert-Stein classification system (Siewert and Stein, 1998). The 8th edition TMN classifies GOJ tumours, as oesophageal tumours with precise staging based upon the point of tumour origin. This classification has been shown to provide better stratification of patients (Talsma et al., 2012) but means that type I&II Siewart-Stein tumours are classified as oesophageal tumours whereas type III tumours fall within the gastric diagnostic criteria. Although this improves stratification of patients, this classification does not necessarily reflect a difference in biology between the two types of adenocarcinoma, and is arguably

an arbitrary division of a diagnostic continuum. Different aetiological factors are likely to shift in the graduation from gastric adenocarcinomas through GOJ adenocarcinomas to oesophageal adenocarcinomas and therefore the current classification system may not be helpful in defining differences in tumour biology.

1.3.2.5 Current understanding of GOA

Trials of GOA treatment have frequently grouped together the different subtypes. Palliative EOX chemotherapy has shown benefit in a cohort of patients with advanced gastro-oesophageal carcinomas, including adenocarcinomas and squamous cancers (Cunningham et al., 2008), whilst, perioperative chemotherapy has been shown to be beneficial in improving survival in locally advanced adenocarcinomas of the stomach, GOJ and lower oesophagus (Cunningham et al., 2006). A clinical trial of preoperative radiotherapy has shown benefit in oesophageal and GOJ adenocarcinomas (van Hagen et al., 2012) and more recently a phase 3 trial involving the use of neoadjuvant FLOT (5-FU, leucovorin, oxaliplatin and docetaxel) has shown a survival advantage over ECX chemotherapy in patients with gastric cancer and GOJ adenocarcinoma (Al-Batran et al., 2017a). Therefore, although the classification system differentiates between oesophageal and gastric adenocarcinomas and squamous cell cancers, the historically important trials and resultant available chemotherapy regimens overlap this divide.

OAC and intestinal type GC arise from a background of intestinal metaplasia, which results from two contrasting risk factors, those of Barrett's oesophagus and *H.pylori* infection respectively. Both processes induce a form of chronic inflammation that progresses to intestinal metaplasia indicating a potential shared pathway between the two anatomical locations (Hayakawa et al., 2016). Where these two risk factor meet at the GOJ it is frequently difficult to determine whether metaplastic changes are secondary to intestinal metaplasia due to Barretts oesophagus in the lower oesophagus or *H.pylori* infection in the gastric cardia (Chandrasoma et al., 2001). Interestingly genetic analysis shows metaplastic cells in Barrett's oesophagus are not related to the surrounding squamous cells (Paulson et

al., 2006) suggesting that they did not originate from squamous progenitor cells but instead may have originated from migrated gastric cardia progenitors (Hayakawa et al., 2016). The described theory implies that there is a common cellular origin of adenocarcinomas arising from Barrett's oesophagus and the stomach (Figure 3). In Figure 3 metaplasia is the replacement of one differentiated cell type by another differentiated cell type in the same tissue. Metaplasia usually occurs due to an environmental stress as a result of the change in expression of a number of transcription factors such as SOX2 and CDX2, however when this stimulus is removed the tissues can return to their normal pattern (Giroux and Rustgi, 2017). Metaplasia is a precursor to dysplasia and Barrett's oesophagus is a type of metaplasia. Dysplasia describes cells that have undergone abnormal development resulting in an expansion of immature cell numbers, and is recognised as a precancerous state which may require treatment. The cells are neoplastic but non-invasive (Sharma and Montgomery, 2013). These cells may show evidence of anisocytosis (unequal cells) poikilocytosis (abnormally shaped cells), hyperchromatism and increased mitotic rate, and cells can be split into grades of dysplasia. In Dysplasia of the oesophagus there are frequent genetic changes seen in genes such as *TP53* and *TP16,* and varying levels of aneuploidy and abnormal methylation (Jankowski et al., 1999).

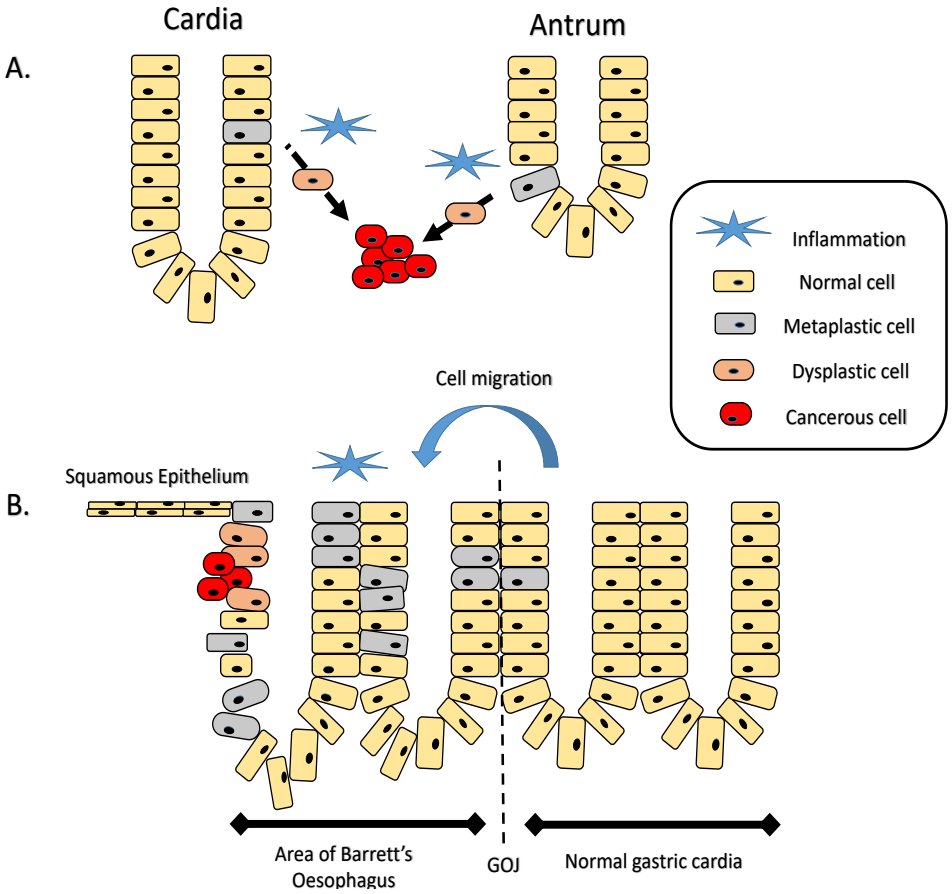


Figure 3. **Potential mechanism of development of GOA.** A. Inflammatory insult in the gastric cardia or antrum leads to the development of metaplastic and dysplastic cells which over time may progress to adenocarcinoma. B. Inflammation at the squamous – gastric epithelial junction leads to expansion of gastric cardia glands into the area of Barrett's oesophagus. Dysplastic cells descending from these migrated cardia cells may then progress to metaplastic cells and then to adenocarcinoma. Figure produced by myself adapted from Hayakawa et al. (2016) and published in Openshaw et al. (2017).

There is therefore evidence for a shared biology of GOAs. Separate environmental factors are believed to cause similar types of inflammation that result in intestinal metaplasia and in addition, there is a potential common cellular origin for adenocarcinomas. These findings taken together with the changing frequency of these cancers and the increasing difficulty in classifying them based upon anatomical location suggests it is not unreasonable to investigate these tumours as one entity.

1.3.3 Genetic Variability of GOA

1.3.3.1 Whole Genome and Whole Exome Sequencing

Dulak et al. (2013) completed an extensive investigation into the genomic alterations associated with OAC and GOJ adenocarcinomas. They undertook WGS of 15 tumour-normal tissue pairs and WES of another 149 pairs. WGS showed a high level of A>C transversions, at 34% of the total mutations. Squamous oesophageal cancers in the same cohort did not share the high A>C transversion rate (Agrawal et al., 2012). A preference for A>C transversions were seen when two adenine base pairs were together resulting in a high frequency of AA>AC transversions (20.2 mutations/Mb) accounting for 29% of total mutations. The rate was most pronounced when the 3' base was guanine causing AAG>ACG transversions with a mutation rate of 49.3 mutations/Mb compared to a rate of 6-17 mutations/Mb for the other 3 bases in the 3' position. The A>C transversion rate also appears to be more common in non-coding regions. A trend was seen for OACs within the tubular oesophagus to have a greater transversion at AA sites than those of the GOJ. This trend led the authors to suggest that these mutations may be attributable to GORD.

Dulak et al. (2013) also screened for significantly mutated genes in OAC & GOJ tumours. Those patients with the highest underlying rate of mutation were all found to have microsatellite instability and appropriate germline mutations. In those patients without a germline mutation, *TP53* and cyclin-dependent kinase inhibitor 2A *(CDKN2A)* were the most significantly mutated genes at 72% and 12% respectively. *TP53* had previously been shown to be important by Chung *et al.* (2007) who analysed exons 4-8 of *TP53* in 40 cases of OAC and showed a mutation rate of 75%. CDKN2A, also known as *p16*, is a tumour suppressor gene that has previously been implicated in the development of OAC (Hardie et al., 2005).

Dulak et al. (2013) highlighted the importance of two genes; engulfment and cell motility 1 *(ELMO1)* and *dedicator of cytokinesis 2 (DOCK2*), mutated at 12% and 6% respectively, as these genes are intracellular mediators *of* ras-related C3 botulinum toxin substrate 1 *(RAC1)* which has been previously implicated in malignant transformation of other cancer types

(Hodis et al., 2012). Three different mutations of *ELMO1* at p.K312 were identified as putative gain-of-function phenotypes. Dulak et al. (2013) also carried out expression studies of mutant *ELMO1* in fibroblast cells which resulted in enhanced cell invasion, thereby supporting the potential role of *ELMO1* in cancer pathogenesis. AT-rich interaction domain *(ARID)1A* and *ARID2* were two other genes with mutation rates ≥5 % which have been implicated as tumour suppressors in a number of cancers including stomach cancer and may therefore be of significance for further investigation. Those *EGFR1* mutations that were identified were not pathogenic whilst those identified in *ERBB2* included mutations in the kinase domain which could be pathologically significant. The *ERBB2* mutations identified in three patients were 2351A>C p.D769Y ×2 and 2327G>T p.G776V. The different mutated genes are summarised in Figure 4.

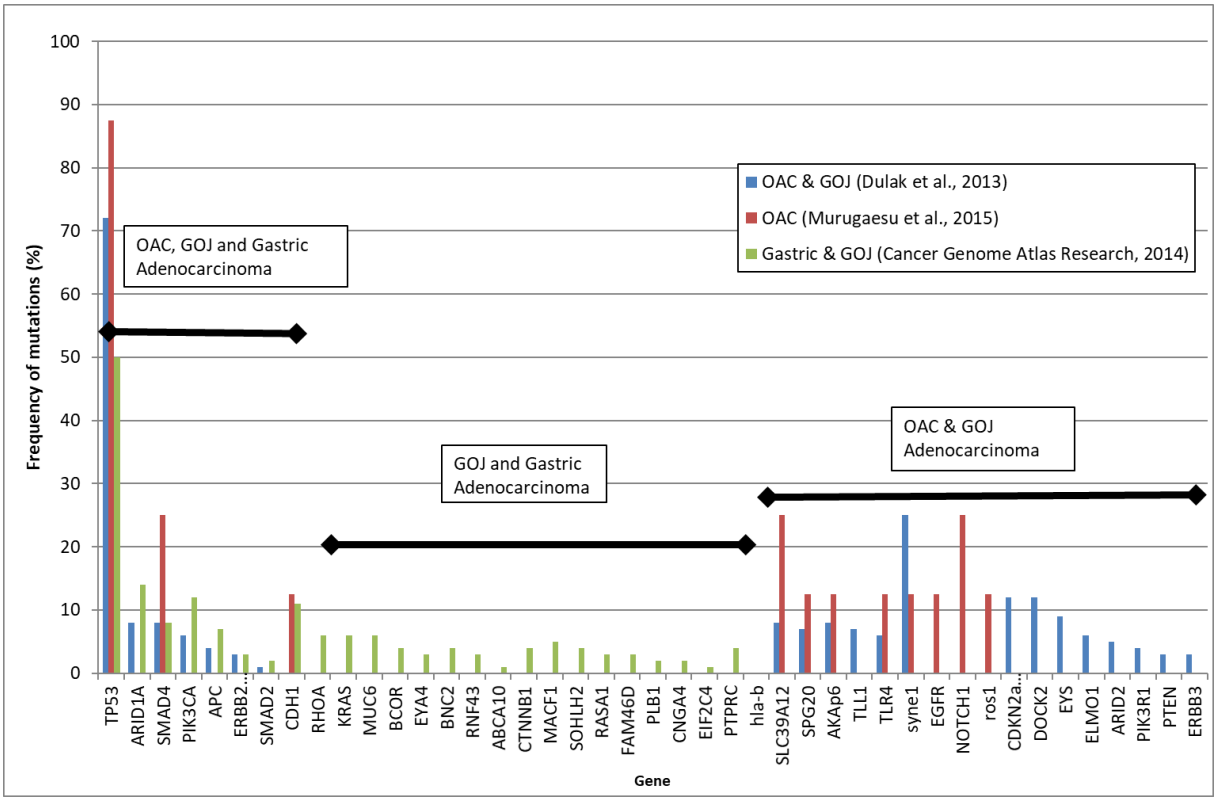


Figure 4- **The Relative Frequency of mutations in Upper GI adenocarcinomas.** Data summarised three studies . To the left are genes that are mutated in all subtypes of GOA.

The mitogen-activated protein kinase pathway (MAPK) was of low significance in the study with no *BRAF* mutations identified and *KRAS* mutations in only 3% of tumours. The Phosphatidylinositol-3-kinase (PI3K) pathway was significantly affected however with a 6% mutation level in *PIK3CA* and other genes including *PTEN and PIK3R1* making a total of 13% of patients carrying a mutation within an important gene in the PI3K pathway.

SYNE1 is a large ~500KB gene that was found to be mutated in up to 25% of patients. However, the q-value, which is a measure of the mutation rate in each gene compared to that which would be expected by chance given a background mutation rate, was less significant than for many other genes, likely due to the large size of the gene. Therefore, mutations in this gene are not likely to be disease driving mutations but more likely to be passenger ones.

A significant degree of chromosomal rearrangement was seen with a median of 172 rearrangements per tumour. This was much higher than the level seen in colorectal cancer tumours, showing that chromosomal instability (CIN) is a significant feature of these tumours.

1.3.3.2 Genetic subtypes of gastric and GOJ adenocarcinomas

The Cancer Genome Atlas Research (2014) Project published data on fresh frozen tissue from 295 patients with gastric and GOJ (19%) adenocarcinomas. They categorised their patients into 4 subtypes based on the analysis of the tumours using six molecular platforms (somatic copy number analysis, WES, DNA methylation profiling, mRNA sequencing, miRNA sequencing and reverse-phase protein array (which is a microarray measuring protein expression levels using antibody dependent binding)). The subtypes were defined as EBV-positive (9%), microsatellite instability (MSI) high (21%), genomically stable (defined as lack of MSI and chromosomal translocations plus *CDH1* and *RHOA* mutations, 20%) and

chromosomal instability (50%). They also observed the same increased A>C transversion rate in AA dinucleotides, especially at AAG as reported by Dulak et al. (2013). If, as suggested, this AAG>ACG transversion is due to an environmental factor such as GORD this, environmental factor must be shared between oesophageal, GOJ and gastric adenocarcinomas. This common environmental factor could be chronic inflammation if the theory of Hayakawa et al. (2016) described in the previous section is true.

The number of mutations varied between the subtypes, with *PIK3CA (55%)* and *BCL6 corepressor (BCOR)* (23%) having high mutation rates in EBV associated tumours. These tumours also showed extensive DNA promoter hypermethylation. The MSI subtype had mutations in major histocompatibility complex class I genes including beta-2-microglobulin and major histocompatibility complex, class I, B *(HLA-B*) which theoretically reduce antigen presentation in a highly mutated cancers, as is seen in melanoma (Bernal et al., 2012). In the genomically stable subtype cadherin 1 *(CDH1) (37%)* mutations were common, a gene which is known to cause hereditary diffuse gastric cancer when mutated in the germline (Blair et al., 2006). *Ras homolog family member A (RHOA)* mutations were also common in the genomically stable subtype (15%), this gene has a role in cell motility and may play a role in the development of diffuse gastric tumours.

Further comprehensive molecular analysis of OAC and GOJ adenocarcinomas has shown that the most common molecular subtype of gastric cancer, CIN, closely resembles that of OAC (Cancer Genome Atlas Research et al., 2017). In this study, there was no clear anatomical point for differentiating GOJ and OAC molecular subtypes and some GOJ tumours were found to be MSI or EBV positive. This shows the similarities between the subtypes OAC, GOJ and gastric adenocarcinomas, and the difficulty defining one from the other. This provides further evidence to treat GOA as a single disease entity.

1.3.3.3 Genome evolution and heterogeneity

Murugaesu et al. (2015) carried out WES on 8 patients and sampled over 40 tumour regions both before and after neoadjuvant chemotherapy. All tumours in their study demonstrated spatial and temporal heterogeneity with approximately half (55.6%) of all potentially pathogenic mutations being heterogeneous. This showed that single tumour biopsies would underestimate mutational burden as has been shown in other cancers (Gerlinger et al., 2012, Swanton, 2012). Interestingly those tumours that were more heterogeneous also had a poorer response to neoadjuvant chemotherapy.

TP53 was mutated in 7 out of 8 cancers and the single patient with functional TP53 had amplification of *MDM2* which is a negative regulator of TP53. Other mutations studied in Barrett's oesophagus have shown similar frequencies of mutation in both Barrett's oesophagus and OAC with the exception of SMAD family member 4 *(SMAD4)* which is found exclusively in OAC. This had led to the proposal that mutations in *TP53* and *SMAD4* are important in the evolution of Barrett's oesophagus to OAC although only 25% of patients with OAC have a mutation in *SMAD4*. Using a new endoscopic technique called the 'Cytosponge' Weaver et al. (2014) were able to sample cells from a large area of the oesophagus and use detection of *TP53* mutations as a biomarker for predicting high grade dysplasia. *TP53* mutations were found in none of the patients with Barrett's oesophagus and no dysplasia, and in 86% of patients with Barrett's oesophagus and high grade dysplasia, thereby indicating the central role mutations in *TP53* play in the evolution of this cancer.

The A>C mutation signature described in Dulak et al. (2013) was also seen in the tumours from Murugaesu et al. (2015) including a predilection for AAG>ACG sites. Evolution studies of the tumour regions showed that these A>C mutations occurred as a signature in all tumour regions suggesting that the pattern occurred before the emergence of the most recent common ancestor of all the regions, as would be expected for an environmental insult that may give rise to the cancer.

Evidence for convergent evolution of different tumour areas was detected in two patients. *NOTCH1* was mutated multiple times in one patient with different mutations of the same gene present in different tumour regions. In another patient *N-acetylglucosamine-1-phosphate transferase alpha and beta subunit (GNPTAB)* mutations showed a similar pattern. Murugaesu et al. (2015) speculated that this was due to similar evolutionary pressures on the different tumour regions later in tumour evolution resulting in different mutations in the same gene in separate tumour regions. This is called convergent evolution as the different mutations have similar effects on the gene but have occurred independently of each other, in contrast to truncal mutations where a single mutation arises in an ancestral cell(s) which gives rise to all tumour areas resulting in the same mutation being present in all tumours.

In a separate study (Findlay et al., 2016) published data on WES of 30 paired OAC tumours, pre- and post-neoadjuvant chemotherapy. They showed that the mutational spectrum of *TP53* mutations may change during anti-cancer therapy and that mutations in other genes may appear. The authors suggest this is due to chemotherapy-induced selection pressures on the sub-clones present. This also demonstrates the dynamic nature of the mutational load in these tumours.

1.3.4 Somatic Copy Number Aberrations

SCNAs are commonly found across all cancers including GOAs and may relfect amplification of oncogenes and deletion of tumour suppressors (Beroukhim et al., 2010). Cells from normal tissue are diploid, with two copies of each gene (2n). Amplification is defined as having at least one additional copy of each gene (2n+1) which equates to a copy number of 3 or more, whilst deletion is defined as the loss of at least one copy of a gene (1n) (Griffiths et al., 1999). It is known that GOAs have more amplification of the genome than other tumours such as colorectal cancer (Dulak et al., 2012). The development of WGS and WES has allowed improved resolution of copy number variants and chromosomal rearrangements compared to array based technology and have shown that GOAs have

highly rearranged genomes (Murugaesu et al., 2015, Dulak et al., 2013, Nones et al., 2014). Chromothripsis; the large scale fragmentation and rearrangement of chromosomes in a single catastrophic cellular event (Storchova and Kloosterman, 2016), occurs in up to a third of patients with OAC. These events are thought to be close to the maximum severity of chromosomal damage that can be tolerated by any cell and provide a mechanism for the ability of OAC to arise suddenly in patients with a background of Barrett's oesophagus; by resulting in sudden amplification of oncogenic regions, or silencing of important tumour suppressors. Nones et al. (2014) speculated that the near ubiquitous lack of function p53 and telomere shortening in OAC promote the occurrence of these genomic catastrophes. Their findings also show that amplification of oncogenes is an important mechanism driving OAC proliferation and that the amplification can occur following a single genomic catastrophe.

In terms of amplified genes the Cancer Genome Atlas Research (2014) identified focal amplification of a number of genes including *ERBB2(HER2), KRAS, MYC, EGFR* and *CCNE1,* (Table 1). The locus 9p24.1 was also recurrently amplified and contains Januse kinase-2 (*JAK2)*, a potentially targetable gene (Furqan et al., 2013), and the genes which encode programmed death ligand 1 (PD-L1) and PD-L2, both of which are characterised immunosuppressant proteins currently under intensive investigation (Mahoney et al., 2015, Ansell et al., 2015). Deletions at the loci of known tumour suppressors included *PTEN, SMAD4, CDKN2A and ARID1A*. The genomic data published by Dulak et al. (2013) has also been analysed to look specifically at copy number variations (Wang et al., 2015a) and amplifications that were potentially driving cancer growth were identified (Table 1). Murugaesu et al. (2015) carried out similar SCNA analysis on the 8 patients included in their study (Table 1). Overall these three studies identified a number of genes with recurrent amplification and loss.

Gene	Gain (%) (Wang et al., 2015a) n=149	Loss (%) (Wang et al., 2015a) n=149	Gain % (Cancer Genome Atlas Research, 2014) n=295	Loss % (Cancer Genome Atlas Research, 2014) n=295	Potential Driver in Murugaesu et al. (2015) Yes/No (No. of Patients out of 8) n=8	Potential Driver in (Wang et al., 2015a)	Chromosomal location
TP53	-	13	-	-	-	-	17p13.1 Deletion
EGFR	16	-	11	-	Yes (1)	Yes	7p11.2 Amplification
ERBB2	19	-	17	-	Yes (2)	-	17q12 Amplification
ERBB3	3	-	15	-	-		12q13 Amplification
MET	6	-	6	-	Yes (1)	Yes	7q31.2 Amplification
KRAS	21	-	17* KRAS/NRAS	-	Yes (1)	Yes	12p12.1 Amplification
PIK3R1	-	15	-	3	-	-	5q13.1 Deletion
PIK3CA	3	-	24	-	Yes (1)	-	3q26.32 Amplification
SMAD2	-	29	-	-	-	-	18q21.1 Deletion
SMAD7	-	28	-	-	-	-	18q21.1 Deletion
SMAD4	-	34	-	-	Yes (8)	-	18q21.2 Deletion
ARIDA1A	-	10	-	-	Yes (5)	Yes	1p36.11 Deletion
CDKN2A	-	32	-	-	Yes (4)	Yes	9p21.3 Deletion
CCNE1	12	-	-	-	Yes (1)	-	19q12 Amplification
CDK6	17	-	-	-	Yes (4)	Yes	7q21.2 Amplification
CCND1	10	-	-	-	Yes (5)	-	11q13.3 Amplification
APC	-	15	-	-	-	Yes	5q14.3 Deletion
ARID2	-	2	-	-		Yes	12q12 Deletion

Table 1 – **Recurrently deleted / amplified regions containing putative driver genes in patients with GOA**.

1.3.5 Summary of GOA genetics

The analyses of the GOA genetics have shown that their genomes are highly complex and characterised by both mutations and SCNAs. A number of oncogenes are highly mutated across all anatomical sites of origin of GOA (OAC, GOJ and gastric adenocarcinoma Figure 4) including *TP53, SMAD4, PIK3CA, CDKN2A,* and these mutations show both spatial and temporal heterogeneity. Chromosomal instability appears to be more prominent in GOA and resultant amplification of genes and SCNAs play an important role tumourogenesis.

Amplified genes (Table 1) include *ERBB2*, which is targetable by trastuzumab (Boku, 2014) and other amplified genes (*KRAS, EGFR* and *PIK3CA)* are again shared across the anatomical types of GOA. These findings, together with the finding that the most common molecular subtypes of OAC and GOJ adenocarcinomas resemble those of the most common gastric cancers suggest GOA can be treated as a single entity. Understanding these genetic changes is important as many may be suitable for non-invasive monitoring using circulating cell free DNA technology as discussed below.

1.3.6 GOA and total cfDNA

In contrast to breast cancer relatively little research on total cfDNA has been completed in patients with GOA. The levels of total cfDNA have been shown to have utility in gastric cancer as a biomarker to differentiate unaffected patients from those with early cancer and may even predict tumour burden (Kim et al., 2014) but has not shown to be useful in other studies (Hamakawa et al., 2014).

1.3.6.1 Somatic mutations and total cfDNA

Hamakawa et al. (2014) screened exons 4-10 of *TP53* in total cfDNA to monitor disease progression and response to surgery. 10 of 42 patients harboured *TP53* mutations by this approach. 6 of the 10 patients had a pre-surgical total cfDNA sample of which 3 had detectable *TP53* mutations in the plasma at baseline. In these 3 patients levels of ctDNA were correlated with progression of disease whereas total cfDNA levels were not. CtDNA remained detectable in these 3 patients post surgery, all of whom relapsed, indicating the potential of this technique to detect minimal residual disease. Outcomes for the other patients were not reported. This study had limited detection of *TP53* mutations to investigate the effectiveness of ctDNA to monitor disease, with only 23.8% of patients harbouring a *TP53* mutation in contrast to the 50-87.5% mutation rate reported elsewhere (Murugaesu et al., 2015, Dulak et al., 2013, Cancer Genome Atlas Research, 2014).

In a recent study on the development of a multi-analyte blood test, Cohen et al. (2018) included 81 patients with GOA in their study of 1005 patients with stage I-III cancers. 64 of the 81 patients had a detectable TP53 (83%) in their plasma, with concentration ranging from 707-0.1 copies/ml and a median of just 1.5 copies/ml. The next most frequently detected mutated gene was *KRAS* with 3 patients (3.6%) harbouring a mutation in their total cfDNA 1.7-0.4 copies/ml. This study did not involve follow-up bloods.

These papers demonstrate the feasibility to detect ctDNA in GOA and the ability to detect minimal residual disease in gastric cancer. Given the similarity between the genetics of gastric cancer and GOA as a whole (demonstrated in section 1.3.3), these findings may be true for all of GOA, and can be further investigated.

1.3.6.2 SCNAs and total cfDNA

GOA shows chromosomal instability and demonstrates amplification and deletion of many genes (see section 1.3.4). Trastuzumab therapy is effective in advanced or metastatic gastric carcinoma where overexpression of HER2 has been confirmed (Bang et al., 2010). HER2 status is determined using histopathology and FISH using the same method as for breast cancer (see section 1.2.4.2, page 30). Heterogeneous HER2 expression has been demonstrated in up to a third of patients with gastric cancer and it is not uncommon for there to be discordant HER2 expression in the primary and metastatic tumours from the same patient (Cho et al., 2013).

Investigation of the utility of total cfDNA to determine HER2 status was carried out by Shoda et al. (2014). In this study qPCR was used to assess *ERBB2* amplification using the standard unamplified reference gene *RPPH1,* which was shown to be superior to *TAOK1* and *EFTUD2*. *ERBB2*:*RPPH1* ratio of tumour tissue was concordant in 37 out of 48 patients (77.1%) with routinely determined HER2 expression status. *ERBB2*:*RPPH1* ratio was highly concordant with FISH analysis. Since samples for routine HER2 status determination and qPCR were not taken from the same place, the discordant results may reflect a difference in sensitivity and

specificity between the two techniques or be a reflection of the high intratumoural heterogeneity of HER2. For total cfDNA analysis a sensitivity of 0.667 and specificity of 1.000 was found for the *ERBB2:RPPH1* ratio when correlated with HER2 status. Of 12 HER2 positive patients, 7 showed *ERBB2* amplification and these 7 patients demonstrated effective response to trastuzumab, whilst tissue IHC/FISH was not predictive of treatment response suggesting qPCR may be a more effective way of determining response to HER2 targeted therapy. Using the same reference gene ddPCR was shown to be both more sensitive and specific than qPCR and three patients with raised *ERBB2* ratio post surgery showed early relapse, suggesting detection of minimal residual disease (Shoda et al., 2017). As with the studies in Breast cancer (Murtaza et al., 2015, Page et al., 2011) some GOA patients acquire ERBB2 amplification over time, with 53.8% of patients that had HER2 negative primary tumours being *ERBB2* amplified at relapse in Shoda et al. (2017). These studies demonstrate the feasibility of longitudinal monitoring of *ERBB2* status in progressive and relapsing patients. This would be of particular benefit in those patients who have not had a recent biopsy or in those patients where one is not possible.

Given that it is possible to detect and monitor *ERBB2* amplification, other genes that are known to be amplified such as *EGFR, ERBB3, KRAS, CCNE1, CDK6* and *CCND1* could also be detected using qPCR and ddPCR.

1.3.7 Conclusion

There is increasing evidence that GOA can be treated as a single disease entity with common genetic aberrations, overlapping anatomical origins and varying aetiological causes. However, despite our improved understanding, GOA continues to have a dismal overall survival and lack of a reliable blood based biomarker. The detection of ctDNA via SNV and SCNA analysis, indicates the value of ctDNA as a new biomarker for GOA which may improve our understanding of the disease and our ability to monitor anti-cancer therapies. Investigation of the utility of ctDNA in GOA is therefore a promising area of research.

1.4 Hypothesis, Aims and objective

Hypothesis

I hypothesised that ctDNA would be detectable in patients with both breast cancer and GOA, providing clinically useful information regarding disease monitoring and detection of minimal residual disease. More specifically, that I would be able to use pre-existing techniques in the lab to monitor palliative breast cancer patients and predict relapse in neoadjuvant patients. Using the same technologies I hypothesised I could develop a sequencing platform for GOA tumours and detect ctDNA in the plasma of patients. The ctDNA should also be able to monitor disease and provide evidence of minimal residual disease.

Finally WES of ctDNA shoul be able to provide enchanced genetic information on the monitoring of a single patient with advanced breast cancer over multiple cfDNA samples.

The work was split into three projects

1. Evaluation of longitudinal total cfDNA analysis by targeted NGS and ddPCR in patients undergoing palliative and curative therapy for breast cancer.
2. Evaluation of WES analysis of tumour DNA and longitudinal cfDNA samples in a single metastatic breast cancer patient.
3. Evaluation of longitudinal total cfDNA analysis by targeted NGS and ddPCR in patients undergoing palliative and curative therapy for GOA.

The aims of the thesis were

1. To develop the translational use of ctDNA patients with breast cancer
2. Demonstrate the ability to carry out WES analysis of ctDNA in a single patient with breast cancer

3. Detect ctDNA in patients with GOA and investigate the potential translational use of ctDNA in this cancer type.

The objectives were to;

1. Investigate the ability to use ctDNA to monitor 13 patients with metastatic breast cancer using the in-house developed AcquiRes NGS panel and lab / commercially available ddPCR assays.
2. Investigate the use of the Oncomine™ cfDNA Breast panel to detect ctDNA in 10 breast cancer patients undergoing curative treatment.
3. Investigate WES analysis of ctDNA in a single patient with breast cancer and analyse the suitability of this technology to interrogate tumour genetics by using ctDNA analysis alone.
4. Detect ctDNA in 44 patients with GOA using targeted NGS and ddPCR / qPCR, and investigate the ability to monitor GOA using ctDNA.
5. Investigate the ability of ctDNA to detect disease relapse in 23 patients with GOA undergoing treatment with curative intent.

2 Materials and Methods

2.1 Study ethics and patient consent

All patients were consented for storage of blood and tissue in the University of Leicester Cancer Research Biobank

- Research & Development Number – UHL11274,
- REC – 13/EM/0196 (East Midlands Local Research Committee).

Patients were informed of the nature of the cfDNA project and of future potential uses of their samples through the biobank as per the consent form (Appendix 7.1). Patients were consented for sampling of past and future tissue samples taken as part of their routine care and for initial and future blood samples of a maximum of 25mls. Patients were free to withdraw at any time point. Patients were informed that there was no advantage to themselves for taking part in the study, apart from the fact that they would know they were helping advance research into their type of cancer that may help other similar patients in the future.

2.2 Patient clinical data and timeline generation

Patient clinical data was collected from patient notes and letters, CT scan reports and iLab pathology reports. iLab is the computer programme used to store pathology data at the university hospitals of Leicester (manufacturer iSoft plc.). Data was recorded using Microsoft access, and timelines generated using Microsoft office timeline for PowerPoint.

2.3 Blood Processing

Up to 25ml venous blood was taken into EDTA tubes and stored immediately on ice for up to 2 hours. An initial centrifugation step of 1000g at 4^{o}C for 10 minutes separated the blood into red cells, buffy coat (lymphocytes) and plasma. The plasma was manually separated and subjected to a further centrifugation of 2000g at 4^{o}C for 10 minutes to remove

remaining cellular debris (Page et al., 2006). Plasma was stored at -80 ^{o}C. For each blood sample a buffy coat (lymphocyte) sample and a red blood cell sample was also stored at -80 ^{o}C.

2.4 Total cfDNA Extraction

Total cfDNA was extracted from 1-3mL of plasma using the QIAamp® (Qiagen) circulating nucleic acid kit, according to manufacturersa guidelines, which had previously been shown to give optimal extraction (Page et al., 2013). In brief, the samples were digested with proteinase K at 60°C for 30 minutes before being passed through a QIAamp® Mini column under vacuum generated negative pressure. Plasma DNA binds to the silica-based membrane whilst contaminants pass through. 3 wash steps removed other contaminants such as cations and proteins which may interfere with subsequent PCR reactions or cause DNA degradation. Total cfDNA was eluted with 100µL of Low TE and stored at 4°C for immediate use and quantification or at -20°C for storage exceeding 6 weeks.

2.5 Freezing of Fresh Tissue

Tissue from selected patients undergoing resection of GOA was obtained from the surgical team within 1 hour of surgery. Tumour was dissected into 2-5µm fragments and fixed onto cork with optimal cutting temperature compound. The tissue was frozen by immersing for 3 minutes in a beaker containing 25ml of iospentane which had been pre-cooled for 5 minutes in liquid nitrogen. Each sample was then stored in the biobank at -80 ^{o}C.

2.6 DNA Extraction from Frozen Tissue

Genomic DNA was extracted from frozen tissue using the QIAamp® DNA mini kit. Initially a H&E slide was prepared for analysis by histopathology to confirm presence of tumour cells. Frozen tissue containing the highest percentage of tumour was cut into sections of 5x5µm by university histology staff and placed into a 1.5ml eppendorf containing buffer ATL

(Qiagen). It was then extracted according to manufacturer's guidelines. This technique used the same DNA binding and purification technology as described in section 2.4. DNA was eluted with 100µL of elution buffer AE (Qiagen) and subsequently quantified using the high sensitivity Qubit® Fluorimeter 2.0 kit. DNA was stored at 4^{o}C until use.

2.7 Obtaining and preparing FFPE Tissue

FFPE samples including initial biopsies, curative resections, palliative resections and relapse biopsies were obtained for each patient. Tumour areas were identified by a consultant histopathologist (for GOA Dr C. Richards, for breast cancer Dr D. Moore) on H&E stained slides from each tumour block. 2mm core biopsies were taken from each sample ensuring that sufficient tissue remained for further histopathological diagnositic procedures to take place. Where possible multiple samples (two tumour areas and 1 lymph node for surgical specimens) were taken from each resected specimen to gain a better representation of tumour heterogeneity. High resolution photographs were taken of all H&E slides (example Figure 5) and the site of each core biopsy was recorded digitally on the secure university R:Drive. Areas of >80% tumour content were delineated by a qualified pathologist, and samples taken within this region as described as above.

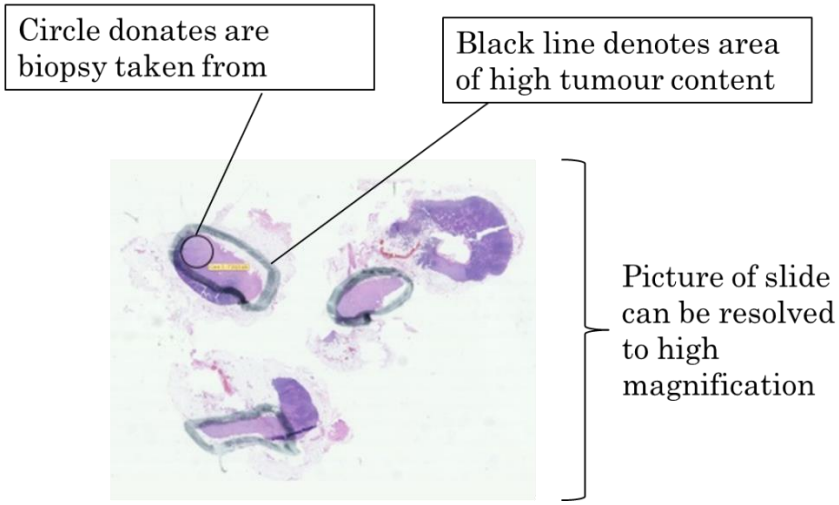


Figure 5 – **Representative H&E specimen.** Areas of high tumour content were delineated by pathologist.

2.8 Extraction of DNA from FFPE tissue using the GeneRead™ kit

DNA from FFPE cores were extracted using the GeneRead™ DNA FFPE kit (Qiagen) according to manufacturer's instructions. The extraction uses the same DNA binding and purification technology as described in section 2.4. The kit involves an additional Uracil-DNA glycosylase (UNG) enzyme incubation step unique to the GeneRead™ kit. Formalin fixation and ageing can result in cytosine deamination to deoxyuracil (Williams et al., 1999) which leads to a cytosine to thymine (C-T) conversion in sequencing reactions as deoxyuracil pairs with adenine. The UNG incubation reverses this alteration, thereby preventing false-positive mutation calling during sequencing (Do et al., 2013). DNA was eluted from the membrane twice in 30µl of Elution Buffer ATE (Qiagen) to a total of 60µl. DNA was quantified using the high sensitivity Qubit® Fluorimeter 2.0 (Invitrogen) kit and stored at 4^{o}C.

2.9 Extraction of germline DNA from white blood cells in the buffy coat

200µl of buffy coat was used for each extraction. The QIAamp® DNA Blood Mini Kit was used and followed as per the manufacturer's protocol (Page et al., 2006). This technique used silica membrane-based technology to purify genomic DNA similar to the total cfDNA extraction method 2.4. Samples were initially incubated with protease for 10 minutes at 56°C. DNA was precipitated with ethanol before being applied to the spin column where DNA binds the silica membrane. Three wash stages were completed as per protocol. DNA was eluted in 200µl of buffer AE (Qiagen) and stored at 4°C. The DNA was quantified using the high sensitivity Qubit® Fluorimeter 2.0 (Invitrogen) kit.

2.10 Taqman® qPCR

In brief the Taqman® technology uses PCR to amplify DNA in the region of interest. A DNA probe is designed, complementary to a sequence within the amplified region. The probe has a fluorophore e.g. FAM (6-carboxyfluorescein) covalently attached to the 5'-end and a quencher e.g. TAMRA at the 3'-end. This construct will not fluoresce whilst the probe is intact. The Taqman® technique employs the 5' DNAse activity of the AmpliTaq DNA

polymerase to degrade the probe, when it is bound in the region of PCR amplification. The degradation of the probe causes dissociation of fluorophore and quencher allowing fluorescence of the fluorophore to be detected (Figure 6). The relative intensity of fluorescence is therefore relative to the amount of PCR product (Holland et al., 1991) allowing real time detection of PCR products.

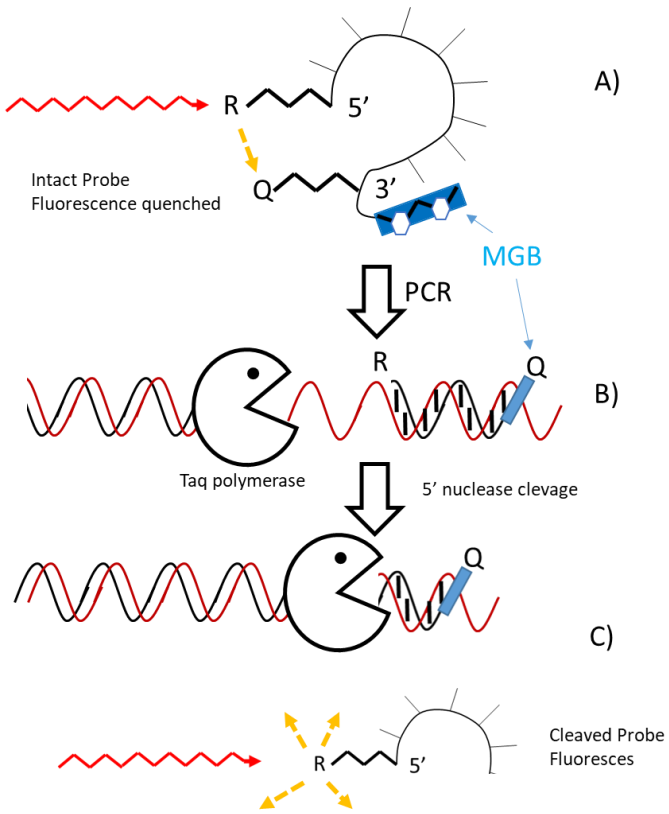


Figure 6 - **Diagrammatic Representation of the Taqman® technique** Showing A) The intact probe, B) Probe binding before degradation by Taq DNA polymerase, C) Cleaved probe. Adapted from Kutyavin et al. (2000).

2.10.1 qPCR quantification of total cfDNA

A previously designed in house locus specific real time qPCR assay, targeting a single copy locus in an intron of the *GAPDH* gene was used to quantify total cfDNA on a StepOnePlus™ machine. The assay contained primers for amplification of a 95bp region of *GAPDH*

(Appendix 7) and a FAM-labelled minor groove binder (MGB) Taqman® probe complimentary to a 20bp sequence within the 95bp region (Page et al., 2011).

Total cfDNA was quantified by real-time qPCR, using absolute quantitation (AQ) against a standard curve of a known dilution of genomic DNA (Roche), each analysed in triplicate. Multiple 1:2 dilutions were used to generate a seven-point standard dilution curve against which the concentration of the total cfDNA could be calculated. The standard curve covered a range of 1.39ng/µL – 21.7pg/µL (5ng-0.078ng per well). The StepOnePlus™ compared fluorescence of the total cfDNA samples to that of the standard curve allowing the concentration of the total cfDNA to be determined.

The *GAPDH* primer/probe mix was combined with TaqMan Fast Universal PCR Master Mix (2X) no AmpErase UNG and 3.6µl of DNA for a total reaction volume of 10µL. All reactions, were run in triplicate on a MicroAmp Fast Optical 96-well reaction plate. The plate was sealed with MicroAmp™ Optical adhesive film and run on the StepOnePlus™ Real-Time PCR System (Table 2).

Stage	Temp/^{o}C	Time	Number of cycles
Denaturation	95	20 sec	
Denaturation	95	1 sec	50 cycles
Annealing	60	20 sec	

Table 2- **PCR conditions for *GAPDH* PCR and DNA quantification**

The above experiment, yielded the concentration of the eluted DNA in ng/µl. The concentration of DNA in copies per ml was calculated as follows:

$$\frac{[\text{Eluted DNA solution in ng/µl}] \times \text{volume DNA eluted in (µl)}}{\text{Volume of plasma used for extraction (mls)}} = \text{ng/ml of plasma}$$

Assuming that the weight of 1 copy of the genome is 3.3 pg

$$(1000/3.3) \times \text{ng/ml} = \text{copies/ml of plasma}$$

2.10.2 Gene amplification detection by qPCR

Detection of gene amplification is possible using relative quantitation (RQ) real-time qPCR by creating a ratio of the cycle threshold (Ct) values for a reference (non-amplified) gene against the Ct value of the gene of interest. The Ct threshold is the PCR cycle where the concentration of the DNA target starts to be detectable above the background noise of the instrument.

The difference between the Ct value can be converted to a copy number value as follows:

$$2^{(\text{Ct gene of interest} - \text{Ct of reference gene}) \times 2} = \text{copy number}$$

Because the primers for the amplicon in the gene of interest and the reference gene will not have identical efficiency and may differ in concentration slightly from experiment to experiment an internal calibration was implemented. Both set of reactions were run against human genomic DNA (hgDNA) and the difference in Ct between the two calculated.

Therefore if (Ct gene of interest – Ct of reference gene) is termed 'dCt'

$$2^{(dCt^{sample} - dCt^{hgDNA}) \times 2} = \text{Calibrated Copy number}$$

Reference genes included *GAPDH, RPPH1* and *CNTNAP1*, the sequences of which are shown in Appendix 7 using the same PCR conditions shown in Table 2.

2.11 DNA quantification and quality assessment by Qubit and microfluidic separation.

2.11.1 Quantification by Qubit 2.0 fluorometer

The Invitrogen Qubit assay uses fluorometry for DNA quantitation. The dye in the Qubit kit greatly increases in fluorescence when it binds to DNA. The amount of fluorescence signal from the dye is directly proportional to the concentration of DNA in the solution, allowing determination of DNA concentration using a Qubit fluorometer. There are kit two ranges;

1. broad range kit, designed to measure DNA from 100 pg/µl–1,000 ng/µl and the
2. high sensitivity kit for concentrations from 10 pg/µl–100 ng/µl.

A master mix was prepared off 199µl of buffer and 1µl of dye per sample. 199µl of the master mix and 1µl of sample DNA was added to each reaction and incubated for 2 minutes. Each kit contains two pre diluted standards; 10 µl of each standard was used to calibrate the Qubit fluorometer 2.0 before quantification of each sample in the fluorometer.

2.11.2 Quantification by Agilent 2100 Bioanalyzer

The Agilent Bioanalyzer was used to profile the range of fragment sizes and to quantify some DNA libraries. A summary of its characteristics is shown below.

The Agilent high sensitivity kit allowed separation and sizing of total cfDNA and DNA libraries of 50-7000bp at concentrations down to 100pg/µl and consisted of a 12 well chip with etched microchannels. The chip and microchannels were filled with a gel polymer and fluorescent dye during chip preparation according to manufacturer's guidelines, and 1µl of sample DNA was loaded into each well. One of the 12 wells in each chip was loaded with a DNA ladder and every sample was also loaded with a lower (35bp) and upper (10380bp) DNA marker. The DNA was driven by a voltage gradient, through the gel, allowing separation of fragments dependent on size. Each sample was analysed sequentially across a single detector to determine DNA fragment concentration. The results were used to assess

DNA quality by assessing the relative amount of different sized fragments and overall DNA quantification.

2.11.3 Quantification by Agilent 4200 TapeStation

The Agilent TapeStation allowed separation of fragmented DNA and quantification in a similar way to the Bioanalyzer. It used pre-prepared gel 'screentapes' to run 1-2 µl of sample on an automated TapeStation machine. The TapeStation could be loaded with high sensitivity 1000 screentapes that could quantify DNA libraries in a similar way to the Bioanalyzer or genomic screentapes for quantification of FFPE, fresh frozen tissue and buffy coat DNA.

2.12 Droplet Digital PCR

Droplet Digital PCR (ddPCR) is a specific technique using Taqman® technology, developed to allow reliable low level detection of SNVs and SCNAs. ddPCR has two steps, firstly the preparation of the PCR mixture and secondly the generation of PCR droplets within an oil emulsion. PCR takes place within each droplet in isolation from the other droplets, allowing thousands of independent reactions to take place. Each droplet can then be analysed according to its fluorescence to determine the DNA template contained within it.

2.12.1 Detection of SNV

SNV detection required the addition of two FRP (forward, reverse probe) mixes to each reaction (Table 3). Both FRP mixes had the same primers. The mutant FRP mix contained a FAM labelled probe complimentary to the mutant allele whilst the wildtype FRP mix contained a VICor HEX labelled probe complimentary to the wildtype sequence. VIC and HEX are fluorophores structurally similar to FAM described in section 2.10, which are released during hydrolysis probe degradation and fluoresce at different wavelengths to FAM. Each reaction contained 5-20ng of DNA.

Reagent	Vol ul	PCR conditions		
ddPCR Super Mix for probes (no dUTP)	11	Enzyme Activation	95°C 10 minutes	
F/R primer and probe mix for Mutation FAM	1.1	Denauration	94°C 30 seconds	40 cycles
F/R primer and probe mix for WT VIC/HEX	1.1	Annealing/Extension	55-65°C for 1 min	
Template DNA	≤8.8	Enzyme deactivation	98°C 10 minutes	
Nuclease free water	To 22	Hold	4°C	

Table 3 - **ddPCR reagent mix and PCR conditions for SNV analysis. F**rom BioRad manual available at http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6407.pdf.

Reaction mixtures were transferred to DG8 ddPCR cartridges and loaded with 70µl of ddPCR droplet generation oil for probes using a QX200™ Droplet Generator (BioRad). At the DNA concentrations used, on average only a single template DNA molecule is incorporated into each droplet. The resultant droplet suspensions were transferred to a 96 well plate before PCR. PCR conditions were as for Table 3 with extension temperature varying depending on the assay used. The ramp rate of the PCR reaction was set to 2°C per second which aids with binding of the complementary probe over the mismatched one. As the droplets undergo PCR the droplets containing mutant DNA template would degrade only FAM probes resulting in a droplet that fluoresced at a different wavelength to those containing wildtype DNA template.

The droplets were read with a QX200™ Droplet Reader (BioRad) using the QuantaSoft™ software using the rare event detection settings for detection of discreet events (Manual available at http://www.bio-rad.com/webroot/web/pdf/lsr/literature/10031906.pdf). The majority of droplets fluoresced at the wavelengths specific to a single probe allowing droplets to be defined as WT or mutant DNA containing. Thresholds for calling droplets were set manually to show a clear separation between mutant or wildtype droplets and negative droplets. Samples containing >3 mutant droplets with generation of >10,000 total droplets were classified as positive. Double positive droplets which fluoresced at both wavelengths, and by chance contained both WT and mutant DNA, were excluded from analysis. By this method a level of detection of 0.3-0.05%% mutant DNA was possible

dependent on amount of DNA loaded, for example if 1,000 positive droplets are obtained in a reaction the lowest level of detection is 0.3%, for 3000 it is 0.1%.

2.12.2 Detection of SCNA

SCNA detection was also possible using ddPCR. SCNA assays contained FRP mixes against two separate genes. The FRP mix containing the FAM probe was designed against a region in the gene of interest, whilst the FRP mix containing the VIC/HEX probe was designed against a region in a gene that is rarely amplified, therefore serving as a control.

PCR setup and droplet generation proceed as per section 2.12.1. At the point of droplet reading the Copy Number Variation settings were used on the QuantaSoft™ software which automatically generates a ratio between the target droplets and control droplets, thereby generating a copy number value.

2.12.3 Design of ddPCR assays

Where possible commercially produced ddPCR assays were used from BioRad. However, a number of variants had no commercially available assay and were therefore designed, validated and optimised as follows. All designed assays are shown in (Table 4).

Gene	Location	Mut.	Forward	Reverse	WT Probe	MUT Probe
TP53	Chr17:7578479G>A	P151S	TGTGCTGTGACTG CTTGTAGA	GCCAAGACCTGCC CTGTG	TG GGT TGA TTC CAC ACC	TG GGT TGA TTC CAC ATC
TP53	chr17:7579320TCA >T	V122Df s26X	CC AGG CAT TGA AGT CTC ATG	TAC CAG GGC AGC TAC GGT TT	TCACAGACTTGGC TGTCCCA	TCAGACTTGGCT GTCCCAGA
TP53	chr17:7579349A>C	F113C	GCCAGGCATTGAA GTCTCA	TACCAGGGCAGCT ACGGTTT	CCAGAATGCAAGA AGCCCAG	CCAGAATGCAA GCAGCCCAG
TP53	chr17:7577127C>A	E271X	AGATTCTCTTCCTC TGTGCG	CTTTTCCTATCCTG AGTAGTGGT	CACCTCAAAGCTG TTCCGTC	CACCTAAAAGCT GTTCCGTC
TP53	chr17:7579312C>A	T125T	GCCAGGCATTGAA GTCTCAT	GTTTCCGTCTGGG CTTCTT	AACTGACCGTGCA AGTCACAGACT	AACTGACAGTGC AAGTCACAGACT
TP53	chr17:7577099C>T	R280K	GCTCCCCTTTCTTG CGGA	GCTTTGAGGTGCG TGTTTGT	GTCTCTCCCAGGA CAGGC	GTCTTTCCCAGG ACAGGC
TP53	chr17:7578206T>G	S215R	AGACCTCAGGCG GCTCATA	GCCCCTCCTCAGC ATCTTATC	CACCACCACACTA TGTCGA	CACCACCACACG ATGTCGA
TP53	chr17:7577560A>C	S241A	GTCTTCCAGTGTG ATGATG	TACCACCATCCAC TACAA	AGGAACTGTTACA CATG	AGGCACTGTTAC ACATG
TP53	chr17:7577141C>A	G266V	ACAAACACGCACC TCAAAGC	ACCTGATTTCCTTA CTGCCTCTT	TCCGTCCCAGTAG ATTACCAC	TCCGTACCAGTA GATTACCAC

Table 4 – **Designed ddPCR assays.** Hg19 coordinates. Mut = mutation.

2.12.3.1 Assay Design

A 200bp region flanking the base of interest was obtained from the UCSC DNA viewer tool (https://genome-euro.ucsc.edu/cgi-bin/). To check the DNA sequence had little or no homology to other regions in the human genome the nucleotide blast tool was used from NCBI (http://blast.ncbi.nlm.nih.gov/). Primer and probes were designed using the Primer3 web-based primer design tools (http://primer3.wi.mit.edu/) aligned to hg19 with the following parameters to optimise the assay for total cfDNA.

i) Primer melting temperature 57-63°C
ii) Max melting temperature difference, 1.0°C
iii) Primer GC content, 40-60%
iv) Product size range, 50-120bp

v) Probe melting temperature 5-10°C higher than that of the primer.

Each designed primer pair was checked using the UCSC *in silico* tool to ensure that a single region of the desired length was amplified. Oligoanalyser 3.1 (https://www.idtdna.com/calc/analyzer) was used to check for primer and probe formation of hairpins and self/heterodimers. A hairpin structure with a Tm less than the reaction temperature and a ΔG of less than -7 kcal/mol for secondary structures was considered acceptable. Where assays could not be designed using Primer3, the primer-probe oligonucleotides were modified from the Primer3 results and / or designed by eye using the same parameters as Primer3. For all designs additional parameters for the primers included no runs of >4 consecutive Gs and no more than 2Gs or Cs in the final 5 nucleotides at the 3' and, for the probes more Cs than Gs, no more than 4Gs and 6 As together and no more than two CC dinucleotides in the middle of the probe.

2.12.3.2 SYBR Green melt curve analysis of primers

Designed Primers were checked for efficiency, specificity, and lack of primer dimer formation by analysis using a SYBR green melt curve. A seven standard curve was prepared to cover the concentration range of total cfDNA; two fold dilution starting at 5ng, each point run in triplicate. Each well contained 5µl Fast SYBR Green Master Mix (2x), 0.6µl forward primer (100pM) and 0.6µl reverse primer (100uM) , 0.2µl sterile H_20 and 3.6µl template DNA. Reaction were run in a 96 well MicroAmp plate sealed with optical adhesive film, on a StepOnePlus™ machine. Cycling conditions were as Table 5 plus a melt curve. Efficiency between 95% and 105% was deemed acceptable parameters for the standard curve. During a melt curve the temperature was increased from 60°C to 95°C with fluorescence readings every 0.3°C to detect the point at which the PCR product dissociation takes place. The SYBR green dye binds dsDNA so as the DNA dissociates the fluorescence is lost. If a single PCR product has been produced all the DNA would dissociate at a specific temperature producing a discrete peak on the melt curve. Primer dimer formation results in multiple PCR products and loss of the discrete peak.

PCR conditions	
95°C 20 seconds	
94°C 3 seconds	
	40 cycles
60°C 30 seconds	
95°C 15 seconds	
Melt Curve	

Table 5 – **SYBR green Melt curve PCR conditions**

2.12.3.3 ddPCR Optimisation

Once primers had been validated the two probes one FAM and one VIC/HEX were ordered. To optimise each assay an annealing temperature gradient was performed. Eight identical reactions were set up using 5ng of genomic DNA as per section 2.12.1 and eight using 5ng of mutation positive tumour DNA. The reactions were run with an eight step temperature gradient above and below the melting temperature of the primers (usually 55-65°C). The optimal annealing temperature was selected that allowed the greatest discrimination between positive and negative droplets whilst avoiding false positive droplets caused by non-specific amplification.

2.13 Targeted Ampliseq™ / Ion Torrent PGM™ semiconductor sequencing

2.13.1 Multiplex and generation of DNA Libraries

Following DNA extraction and quantification, 10-20ng of DNA was used to generate libraries using the Ion AmpliSeq™ Library Kit 2.0 (Life Technologies). The initial multiplex PCR reaction used one of two Ampliseq™ panels tailored to breast cancer and GOA; the AcquiRes breast cancer panel was a previously in-house designed panel against regions in four genes, *ESR1, HER2, TP53* and *PIK3CA* (Table 8) (Guttery et al. 2015) whilst the GOA panel was designed in section 5.4 (Page 157).

Each panel consisted of a pool of primer pairs that amplified 80-120bp regions of interest defined during the design process. The initial multiplex PCR reactions were set up as per Table 6 with a variable number of PCR cycles depending on the type of sample. This was

previously defined for the Acquires panel (Table 6) and defined during the optimisation process for the GOA panel (section 5.5.2, page 162). The multiplex step resulted in an amplicon library enriched for the gene sequences of interest.

Reagent	Volume	PCR Conditions	
5X Ion AmpliSeq™ HiFi Mi	4µl	99°C 2minutes	Acquires Panel Multiplex cycles
2X Ion AmpliSeq™ Primer Pool	10µl	99°C 15 seconds	20 cycles buffy coat & fresh frozen DNA samples
DNA	≤6µl	60°C 2 minutes	25 cycle FFPE DNA samples 28 cycles cfDNA samples
Nuclease free Water	To 20µl total volume	Hold 10°C	

Table 6 – **Multiplex PCR: Reagents and PCR conditions for Ampliseq™ library preparation**

The amplicon libraries were prepared for sequencing following a number of steps as shown below:

1. Partial digestion of the amplicon by addition of 2µl FuPa reagent and incubation as per manufacturers guidelines. The exact method of digestion of FuPa was proprietary to the manufacturer. This step degraded remaining primers and prepared the amplicon library for the next step.
2. P1 adapters and Ion Xpress™ barcodes ligation by addition of 4µl switch solution, 2µl of prepared adapter/barcode solution and 2µl DNA ligase and incubation as per manufacturers guidelines. The P1 adapter DNA sequences ligate to the 5' and 3' end and allow for later amplification of correctly ligated amplicons. A separate 'barcode' DNA sequence was ligated for each DNA library to allow assignment of each DNA molecule to the correct library during sequencing analysis.
3. Library purification using Agencourt AMPure XP™ beads (Beckman Coulter) using two 80% ethanol washes to remove the reagents and buffers from the previous step.
4. Amplification of sequences with the appropriate P1 adapters. This involved bead resuspension in 50µl of Platinum™ PCR supermix HiFi to elute the DNA and prepare them for further amplification using library amplification primer as per Table 7.

Reagent	Volume	PCR Conditions	
Platinum™ PCR Supermix HiFi containing amplicon DNA	50 µl	98°C 2minutes	
Library Amplification Primer	2 µl	98°C 15 seconds 64°C 1 minutes	7-10 cycles depending on DNA sample
		Hold 10°C	

Table 7 – **Regeant and PCR for Library amplification step**

5. AMPure XP™ bead purification. A second purification step with two 80% ethanol washes as per manufacturers guidelines.
6. Library Elution in 50µl of low TE and quantification using TapeStation HSD1000. These libraries were then appropriately diluted with nuclease free water to obtain 100pM libraries.

2.13.2 Template ISP preparation using Ion One Touch™

Between 8 and 12 libraries could be used on each sequencing run depending on the depth of sequencing required. 10ul from each 100pM library were combined to make the final sequencing library before incubation with the Ion sphere™ particles (ISPs) to bind the amplicons from the 100pM libraries to the ISPs. The ISPs were then subjected to an emulsion PCR using the Ion PGM™ Template HiQ View OT2 200 kit on the Ion OneTouch™ 2 system as per manufactures guidelines.

The ISPs were coated in a sequence complimentary to the P1 adapter which binds the amplicons. At 100pM a single DNA molecule should bind to each ISP which was then clonally replicated during the emulsion PCR. If two or more molecules bound, then polyclonal ISPs were formed which could not be sequenced successfully.

Following this step enrichment of the ISPs and removal of the reaction agents and buffers was completed using Dynabeads® MyOne™ Steptavidin C1 Beads on the automated Ion

OneTouch™ ES instrument. Quality control of ISP enrichment was carried out using the Ion Sphere™ Quality Control Kit.

2.13.3 Ion Torrent sequencing of DNA templates using the Personal Genome Machine™

Ion Torrent sequencing uses semiconductor technology to convert changes in pH during sequencing into digital information that can be recorded and analysed by the Torrent Suite™ software. The prepared ISPs were sequenced using the Personal Genome Machine (PGM)™ Sequencing HiQ View 200 Kit. The enriched ISPs were loaded onto an Ion PGM™ 316 semiconductor chip by flowing the solution containing the ISPs across the chip and then centrifuging the chip with the sample in place. The Ion PGM™ 316 chip was then placed in the Ion PGM™ sequencer and a total of 400 flows of dNTPs were undertaken during each sequencing run.

The method by which the PGM™ sequencer works is as follows. The Ion PGM™ chip consists of millions of pH-sensitive wells into which ISPs are placed during centrifugation. During initialization of the PGM™ Sequencer the pH of the solutions used in the sequencing process are brought to a tightly controlled pH. Solutions containing the four nucleotides are loaded into separate sections of the PGM™ to allow control of the sequencing process. Sequence determination is achieved by the PGM™ sequentially flowing each of the four dNTPs over the chip. Each time a base is incorporated a H^+ ion is released inside the well containing the ISP resulting in a change in pH which is recorded by the sequencer (Figure 7). The number of bases incorporated is proportional to the change in pH. Using the change in pH and the sequential cycling of the four dNTP solutions the full sequence of every DNA fragment can be determined.

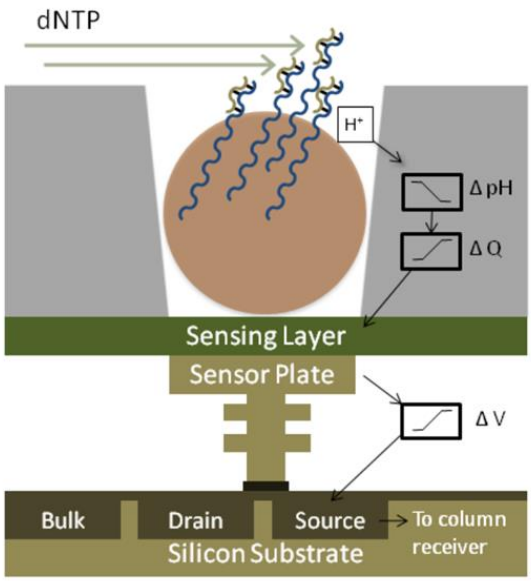


Figure 7 - **Schematic representation of a single well in an Ion Torrent sequencing chip.** The well contains an ISP coated in template DNA. On incorporation of each nucleotide a proton is released changing the pH which is detected in the sensing layer and converted to a change in voltage (ΔV)

2.13.4 Data Analysis

Sequencing data was analysed using the Torrent Suite™ software (version 5.6). Variants were called using the variant caller plug-in which is designed to detect recurrent mutations in each barcoded sample. All called variants were individually analysed using the Integrative Genomics Viewer (IGV) software (version 2.3.82) to ensure they were not due to an artefact of the sequencing process. In particular, called variants that were only present in the last 3 bases of reads were excluded. Variants called in tumour sample but not present in corresponding total cfDNA samples were also analysed to look for low level mutation rates. The significance of each mutation was confirmed using the UCSC Genome Browser, Catalogue of Somatic Mutations in Cancer (COSMIC) and the Database of Single Nucleotide Polymorphisms (dbSNP).

For the AcquiRes panel, variants were reviewed and accepted as true variant if they had:

- Quality score > 20
- Reference reads >30

- Mutant reads >5
- Frequency > 1%
- Strand bias < 0.45
- Reviewed on IGV and not in last 3 bases of read

Potential mutations not fitting these criteria were only accepted as true variants following confirmation with a mutation specific ddPCR assay.

Chromosome	Hg19 Chromosomal Co-ordinate Start Location	End location HG19	Amplicon
chr3	178921449	178921571	PIK3CA_1
chr3	178928068	178928160	PIK3CA_2
chr3	178936052	178936179	PIK3CA_3
chr3	178938796	178938874	PIK3CA_4
chr3	178947747	178947868	PIK3CA_5
chr3	178952017	178952100	PIK3CA_6
chr6	152129003	152129098	ESR1_1
chr6	152129325	152129460	ESR1_2
chr6	152265312	152265434	ESR1_3
chr6	152332762	152332848	ESR1_4
chr6	152415468	152415563	ESR1_5
chr6	152419903	152419999	ESR1_6
chr6	152420050	152420143	ESR1_7
chr17	7573929	7574052	TP53_1
chr17	7577014	7577142	TP53_3
chr17	7577496	7577632	TP53_2
chr17	7578179	7578307	TP53_4
chr17	7578347	7578484	TP53_6
chr17	7578515	7578650	TP53_5
chr17	7579348	7579483	TP53_7
chr17	7579857	7579966	TP53_8
chr17	37868131	37868230	ERBB2_1
chr17	37871757	37871853	ERBB2_2
chr17	37872086	37872187	ERBB2_3
chr17	37879587	37879689	ERBB2_4
chr17	37880182	37880281	ERBB2_5
chr17	37880970	37881064	ERBB2_6
chr17	37881291	37881385	ERBB2_7
chr17	37881561	37881645	ERBB2_8
chr17	37884086	37884180	ERBB2_9

Table 8 - **The Hg19 location of the amplicons covered by the AcquiRes panel.**

2.14 Oncomine™ / Ion Torrent S5 semiconductor sequencing

The Oncomine™ cell-free nucleic acid assays are specifically designed to detect somatic mutations down to a level of 0.1% when starting from 10ng cfDNA template. The Oncomine™ Breast total cfDNA assay version 1 was used to analyse total cfDNA samples from breast cancer patients undergoing neoadjuvant treatment. This panel includes

hotspots in 10 genes: *AKT1, EGFR, ERBB2, ERBB3, ESR1, FBXW7, KRAS, PIK3CA, SF3B1, TP53.* A unique molecular tag is attached to the gene specific primers. This allows the amplicons generated during the multiplex PCR to be grouped into families. Each family is multiplexed from the same original strand of DNA. Therefore, each family should contain the same mutant variant, whilst random errors generated through the sequencing process can be removed during variant calling with specially developed software. The ability to tag each initial DNA molecule also allows more accurate determination of mutations.

We used the Oncomine™ cfDNA breast assay v1. The first step was to amplify targets from the cfDNA.

Reagent	Volume	PCR Conditions	
cfDNA Library PCR Mastermix	15µl	98°C 2 minutes	
cfDNA Oncomine Breast Assay	2µl	98°C 30 seconds 64°C 2 minutes 62°C 2 minutes 60°C 2 minutes 58°C 2 minutes 72°C 30 seconds	} Amplify targets 2 cycles for cfDNA
DNA	≤13µl		
Nuclease free Water	**To 30µl total volume**	Hold 4°C	

Table 9 – **Oncomine™ Target amplification PCR conditions.**

Following the amplification step each sample was purified using AMPure XP™ beads using 80% ethanol. The 30µl of reaction volume was added to 45ul of AMPure XP™ beads and incubated for 5 minutes before carrying out two washes as per section 2.13.1 with 80% ethanol. The final pellet was resuspended in 23µl of low TE.

Purified samples were then amplified with barcoded primers using individual Tag sequencing barcodes to allow sample pooling.

Reagent	Volume	PCR Conditions	
DNA from AMPure XP™ beads purification step	23µl	98°C 2minutes	
cfDNA Library PCR mastermix	25µl	98°C 15 seconds 64°C 15 seconds 72°C 15 seconds	} Barcode Ligation 18 cycles for cfDNA
cfDNA Library Primer P1	1µl		
Tag Sequencing barcode	1µl	72°C 5 minutes	
Total Volume	**50µl**	Hold 4°C	

Table 10 **– Oncomine™ Muliplex PCR with barcoded primers**

Following the amplification step each sample was purified using AMPure XP™ beads using 80% ethanol**.** The 50µl of reaction volume was added to 57.5ul of AMPure XP™ beads and incubated for 5 minutes before carrying out two washes as per section 2.13.1 with 80% ethanol. The final pellet was resuspended in 50µl of low TE.

To size select the barcoded library, 50µl of sample was added to 50µl of AMPure XP™ beads and incubated for 5 minutes before carrying out two washes as per section 2.13.1 with 80% ethanol. The final pellet was resuspended in 30µl of low TE.

Each library was quantified using the Agilent Tapestation HSD1000 kit using values calculated from a designated 100-300bp region. Each library was diluted to 100pmol with low EDTA TE. Using this Oncomine assay, 12 different samples could be pooled and run together on a 530 chip. A temporary library was made by pooling 10ul of each individually barcoded library/sample. The pooled libraries (25ul) were added to the Ion Chef reagent cartridge, after which the IonChef carried out automated templated ISP preparation (2.13.2) and chip loading overnight. This reduces user dependent errors and no QC of templated ISPs are necessary, however there is an increased cost. The following day the chips were loaded into the S5 sequencer which uses the same IonTorrent sequencing technique as the PGM™ (2.13.3). The sequencing metrics, coverage analysis and variants can be viewed using the Torrent Suite Software package.

2.15 NanoString nCounter®

The NanoString nCounter® uses a highly multiplexed probe based system for detecting SCNAs in 87 commonly amplified or deleted genes in cancer (Table 11). 300ng DNA from GOA FFPE and FF samples were analysed using the nanoString nCounter® v2 cancer CNV assay. This was carried out as a fee for service at the UCL NanoString centre (https://www.ucl.ac.uk/biosciences/research/accordion/nanostring).

nCounter® Cancer CN Assay Gene List												
AKT2	BCL2L2	CCNE1	CSMD1	ERBB2	GRB2	KIT	MDM2	MYCL1	PARK2	PTPRD	TERT	YAP1
AKT3	BIRC2	CDK4	DCC	FADD	HMGA2	KRAS	MDM4	MYCN	PAX9	RB1	TP53	YWHAZ
APC	BRCA1	CDK6	DCUN1D1	FGFR1	IGF1R	MAGI3	MELK	NCOA3	PDE4D	REG4	TP73	ZNF217
AR	BRCA2	CDKN1A	DYRK2	FHIT	IRS2	MAP2K4	MET	NF1	PDGFRA	REL	TRAF2	
AURKA	C8orf4	CDKN2A	E2F3	FOXO1	ITGB4	MAP3K5	MITF	NKX2-1	PIK3CA	RPS6KB1	VEGFA	
BBC3	CCND1	CDKN2C	EEF1A2	GAB2	JUN	MAPK7	MYB	NKX2-8	PRKCI	SHH	WHSC1L1	
BCL2L1	CCND2	CRKL	EGFR	GPC5	KDR	MCL1	MYC	ORAOV1	PTEN	SKP2	WT1	

Table 11 – **NanoString nCounter genes included in analysis**

2.15.1 NanoString Technology

The nanoString nCounter® v2 cancer CNV assay has three probes complexes for each target gene, with each probe complex binding a different 100bp sequence within the gene.

For each 100bp target region there are two 50bp probes. The 50bp capture probe is biotin labelled allowing capture of the probe onto streptavidin coated surfaces. The other 50bp probe is covalently linked to a colour coded molecular barcode to form the reporter probe (Figure 2, page 23). Each target sequence will bind a different colour coded probe. Prior to hybridisation the FFPE DNA is fragmented by the restriction enzyme Alu I to allow easy binding to the probes.

Hybridisation of the DNA the probes with the DNA sample of interest takes place in solution. There is an excess of probe to ensure all target DNA is bound. The capture and reporter probes are held together by the target DNA, before the probe complex is elongated and immobilised. The complex is then stretched in one direction by application of an electrical field to allow easy imagining of the barcodes. A digital analyser counts the number of colour

barcodes related to each target. Each of the steps above is automated. As there is no PCR a single count is equivalent to a single molecule of target DNA carrying the sequence of interest, making the assay highly sensitive and reproducible. However as there is no PCR step 300ng of FFPE DNA is needed for each sample, and this means it is not suitable for use in total cfDNA analysis.

2.15.2 NanoString Data Analysis

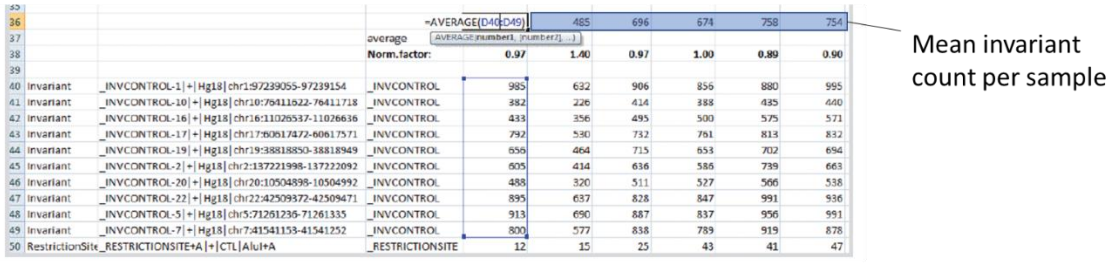
400ng of each sample was sent to the UCL NanoString centre for analysis. The results were returned as a spreadsheet with the sample identifiers and the digital 'counts' recorded for each probe. The nCounter assay has a number of inbuilt controls to ensure that the various steps have taken place correctly.

1. Six positive control probes each targeting a unique DNA sequence added to every sample. The controls are at 0.125, 0.2, 2, 8, 32, and 128fM allowing a standard curve to be calculated for every run. This ensures that the hybridisation reaction, capture and analysis has been completed correctly.
2. Eight negative control probes, for which there is no template DNA, to monitor the background count rate for each run.
3. Ten Invariant probes, designed to target genomic regions predicted not to contain common SNCAs.
4. Four 'restriction' probes. Two probes (A+B) are designed to contain an Alu restriction site to monitor that Alu digest has happened correctly. Two probes (C+D) lack an Alu site and serve as a control that template DNA was present at the time the probes were added and the Alu digest was undertaken.
5. 261 custom SCNA probes designed by nanoString for the 87 genes in the nCounter v2 CNV assay

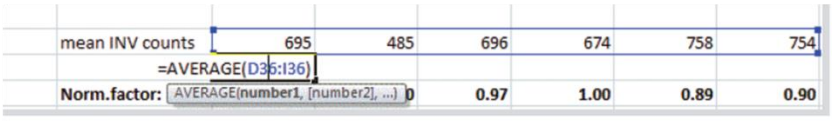
As the nCounter contains probes designed to invariant regions of autosomes. It is assumed that these regions represent 2 chromosomal copies in a vast majority of samples analysed.

Therefore, normalizing data to the counts obtained from these invariant probes corrects for differences in sample to sample genomic DNA input arising from pipetting error or inaccuracies in DNA quantitation. Data was normalised across the run if all samples were of the same type e.g. FFPE or Fresh Frozen, or only between samples of the same type. This was carried out by averaging the number of counts for each invariant probe to get a mean invariant count per sample (Figure 8a). The average of the mean invariant count per sample was then found to get an overall run mean (Figure 8b). The overall mean was dived by the invariant mean count per sample mean to get a normalisation factor (Figure 8c). This factor was used to multiply the counts by to correct for loading errors before calculating the CNV.

a) Generation of mean invariant counts per sample



b) Average of mean invariant counts



c) Generation of normalisation factor

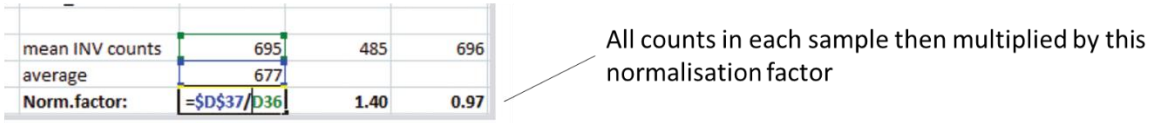


Figure 8 – **Method of normalisation of count values in NanoString.**

To calculate the CNV for each sample the counts for each target were compared to those for a reference sample. As variations in probe efficiency can depend on the quality of the

DNA, a reference sample was used that has been prepared in the same way, ideally DNA from paired normal tissue from the same tissue specimen. If DNA from normal tissue was not available, DNA from tissue was used that had been collected in a similar way i.e. DNA from FFPE was compared to DNA from another normal FFPE sample. DNA from fresh frozen tissue which shows little degradation was compared to DNA from the Coriell Insitute sample NA10854 as a reference for which there is publicly available genotype data (http://www.internationalgenome.org/data-portal/sample/NA10854). CNVs were obtained by dividing the number of counts for the sample of interest by the number of counts for the reference sample to get a ratio. This ratio was then multiplied by 2 to get the CNV.

2.16 Whole Exome DNA Sequencing – Data Pre-processing

2.16.1 Library preparation and WES

WES was carried out in Spain in collaboration with Dr R. Toledo. A summary of the technique used is as follows, and is summarised in (Toledo et al., 2018).

Whole exome sequencing libraries of cfDNA, FFPE and germline / buffy coat (20-40ng) samples were prepared using the ThruPLEX Plasma and DNAseq Kits. Samples were barcoded during eight PCR cycles of template preparation, and 550 ng of each sample was processed through the SureSelectXT Target Enrichment System (Agilent SureSelect V5, ref. 5190-6208, protocol G7530-90000 version B1). Captured targets were subsequently enriched by 11 cycles of PCR with KAPA HiFi HotStart (Kapa Biosystems), with the following primers, which target generic ends of Illumina adapters: AATGATACGGCGACCACCGAGAT and CAAGCAGAAGACGGCATACGAGAT. For sequencing, magnetic bead–purified libraries with similar concentrations of cfDNA and tumour DNA derived libraries, and half the concentration of germline DNA libraries were pooled. Sequencing was carried out on the Illumina HiSeq4000 platform. (Toledo et al., 2018)

2.16.2 Data Pre-processing

Prior to analysis for sequence variants the paired end FASTQ files must be cleaned for technical biases caused by the sequencing process. The genome analysis toolkit (GATK) guidelines were followed (McKenna et al., 2010). The sequences were initially analysed using the programme FASTQC (version 0.11.5) to look for evidence of residual adapter sequences and duplicated reads caused by the PCR steps. File processing and variant calling was carried out by Dr Robert Hastings, bioinfomatician. The BBASH DNA-Seq Next Generation Sequencing Analysis Workshop - 2nd-4th May 2017 – was attended to provide understanding and practical application of each step. The specific steps were as follows

1. Trim illumina adapters using trimmomatic (version 0.36) IlluminaClip (Bolger et al., 2014).
2. Make unmapped binary alignment map (BAM) file
3. Use picard (version 2.7.1) MarkIlluminaAdapters to mark residual adapters
4. Convert unmapped BAM back to FASTQ removing any marked adapter sequencing
5. Convert FASTQ file to aligned sequence alignment map (SAM) file using Burrows-Wheeler Aligner, BWA (version 0.7.15) (Li and Durbin, 2009).
6. Convert mapped SAM file to aligned BAM using picard (version 2.7.1)
7. Mark duplicate sequences using picard (Broad, 2018) and remove them.
8. Base (Quality Score) Recalibration using and GATK (version 3.6) (McKenna et al., 2010)

Steps 1-4 involved removing the Illumina adapters. These are specific sequences of DNA that were added to each sample during sequencing so that DNA from each sample carried a unique barcode. This allowed each sequenced DNA sample to be assigned to the sample from which it came following sequencing. The software from the illumina sequencing removed the majority of the barcode sequences before we received them, however some adapter sequences remained which must be removed before analysis. For the majority of samples, the illumina adapters were not present in the sequences at high levels according

to analysis with the programme FASTQC (version 0.11.5). However, in sample T15-1 there was significant residual adapter sequences and these were removed using a two step process. First the files were processed with trimmomatic (version 0.36) IlluminaClip. Secondly the files were converted from FASTQ to BAM format and processed using the programme picard (version 2.7.1) MarkIlluminaAdapters, which marked any residual adapters. The marked sequences were then removed in the conversion of the BAM file back to a FASTQ file, by first assigning marked sequences a low Phred score (a measure of the quality of the identification of the nucleobases generated by automated DNA sequencing) and then clipping bases marked with the low Phred score. These steps successfully removed any residual adapter sequences as shown for sample T15-1. Any low level residual adapter sequences present in the other samples were also removed via this processing method.

Steps 5-6 involved the alignment of the raw sequencing FASTQ file to the human genome (Hg19) using BWA (version 0.7.15) (Li and Durbin, 2009). This produced an aligned BAM file for each sample sequenced.

Steps 7 used the paired end reads from the aligned BAM to identify reads that are likely to have originated from duplicates of the same original DNA fragments. These occurred because of the nature of the multiple PCR cycles involved in sequencing. All but one of these reads identified as a duplicate was removed using picard before variant calling (Broad, 2018).

Step 8 searched for biases in the Phred quality scores for each base. Phred scores are highly useful in weighing the evidence for or against possible variant alleles during variant calling. Bias can occur due to the chemistry of the library preparation and sequencing or due to defects in the sequencing machinery. Recalibration is a complex statistical process using the data from all the base calls across all the samples to produce a model that identifies any biases. GATK (version 3.6) was used to apply base quality adjustments to the dataset.

2.16.3 SNV and INDEL analysis

Somatic SNVs and INDELS were called with Mutect2, part of the GATK3 v3.6 suite of tools (McKenna et al., 2010). Functional annotation was performed with Variant Effect Predictor (v86), (McLaren et al., 2016) via the vcf2maf program (Inc., 2018), so each variant was mapped to only one of all possible gene transcripts/isoforms that it might affect.

2.16.4 Driver gene identification

TIER 1 and TIER2 drivers were identified using the CGI (Tamborero et al., 2018) for SNVs, and SCNA using the web based analysis function at https://www.cancergenomeinterpreter.org/home. Tier 1 mutations have documented evidence of activity that may drive cancer, as well as evidence that changes to protein activity promotes oncogenic transformation. Tier 2 mutations have strong evidence for playing a role in cancer but this is less comprehensive than for tier 1.

2.16.5 Mutational Signature Analysis

To identify mutational processes that may be driving the tumours, the SNVs were analysed to extract the mutational signatures in the samples using the R package Maftools with the extract signatures function (Mayakonda et al., 2018). We then compared the extracted mutational signatures once decomposed to known signatures derived from Alexandrov et al. (2013) and cosine similarity was calculated to identify best match and the relative contribution in each sample. Large scale genetic analyses of cancers have revealed many mutational signatures across the full range of human malignancies. The current set of mutational signatures is based on The Cancer Genome Atlas and the International Cancer Genome Consortium. There are over 30 different signatures and each signature has a specific pattern of nucleotide substitutions many of which are correlated with specific causes. For example, signature 7 has large numbers of CC>TT dinucleotide mutations caused

by exposure to ultraviolet light and is found in skin cancers, whereas signature 3 has changes associated with failure of double stranded DNA break repair by homologous recombination typical of patients with somatic BRCA mutations (Petljak and Alexandrov, 2016).

2.16.6 Copy number profiling with Sequenza

To determine the somatic Copy Number Aberration profiles of the FFPE and cfDNA-WES samples we applied the allele-specific copy number R package Sequenza v2.1.2 (Favero et al., 2015). The processed final BAM files from the alignment and QC process of tumours and matched lymphocyte were used as input. The estimated purity and ploidy of the tumour was determined (appendix 7.5) and the profiles generated. The segmented regions were then mapped to a cytoband location along with the genes within using Bedtools (Quinlan and Hall, 2010).

3 ctDNA profiling in patients with breast cancer undergoing palliative and curative therapy

3.1 Introduction

ctDNA can act as an alternative to metastatic biopsies for mutational analysis (Rothe et al., 2014) and can be detected in patients undergoing both palliative and curative treatment (Dawson et al., 2013, Rothe et al., 2014, Olsson et al., 2015). ctDNA has a role in tracking tumour specific mutations allowing disease monitoring (Dawson et al., 2013, Page et al., 2017), detection and prediction of resistance to anti-cancer therapies (Guttery et al., 2015, O'Leary et al., 2018, Fribbens et al., 2018) and prediction of relapse in patients treated with curative intent (Garcia-Murillas et al., 2015).

These studies suggest that the majority of patients with metastatic breast cancer have detectable ctDNA that can be used to monitor disease response and progression and that the detection of ctDNA in patients undergoing neoadjuvant chemotherapy may also have important clinical utility. The approach used in this chapter is based upon that used by (Garcia-Murillas et al., 2015), who used genetic profiling of primary tumour DNA by targeted sequencing, to identify high frequency somatic mutations which were then tracked in total cfDNA. This method detected a higher proportion of patients with a trackable mutation in total cfDNA than did a similar study without a tumour DNA screening step (Oshiro et al., 2015). These high frequency 'driver' mutations can be tracked by ddPCR (Garcia-Murillas et al., 2015, Dawson et al., 2013) or by NGS sequencing (Page et al., 2017) and therefore both techniques were utilised in this chapter.

Given the growing body of research, it was of significant interest to apply these techniques to breast patients undergoing routine anti-cancer therapy to investigate the translational

use of ctDNA in both the metastatic and curative intent setting. Therefore, this chapter investigated the use of the liquid biopsy in routine follow up of 35 women with breast cancer, 13 of whom had metastatic disease and 22 of whom received neoadjuvant chemotherapy with curative intent.

3.2 Hypotheses, Aims and Objectives

3.2.1 Hypotheses

The hypotheses to be tested was that ctDNA could be detected in the majority of palliative breast cancer patients. That ctDNA levels would fluctuate with disease response and progression and that a rise in levels may predict failure of therapy. In addition, emergent mutations in *ESR1* may be detected. Finally it was predicted that ctDNA maybe detectable in patients undergoing neoadjuvant chemotherapy, and that these results may predict which patients will relapse.

3.2.2 Aims

The aims of this chapter were to:

1. Compare the findings of tumour DNA and ctDNA NGS analysis using the in-house developed AcquiRes panel, and monitor patients using the serial analysis of ctDNA samples.
2. Determine if ctDNA analysis has a role in patients undergoing neoadjuvant chemotherapy

3.2.3 Objectives

Palliative breast cancer patients:

1. Compare paired tissue DNA / total cfDNA analysis by targeted NGS sequencing using a custom NGS panel.
2. Determine changes in ctDNA over time during anticancer therapy, using a combination of:
    i. Ampliseq NGS total cfDNA analysis using the AcquiRes targeted panel.
    ii. Total cfDNA analysis by ddPCR for patient tumour specific mutations
    iii. Total cfDNA analysis by ddPCR for *ESR1* mutations.

Patients undergoing neoadjuvant treatment for their breast cancer:

1. Assess ctDNA dynamics during neoadjuvant chemotherapy by ultra deep sequencing using a commercial NGS Oncome™ panel.

3.3 Patient Recruitment and characteristics

Recruitment of 13 palliative patients was completed by May 2016. Patients were followed up with repeat blood sampling and collection of clinical history until August 2018 or death. Two patients died before additional bloods could be collected. Patient recruitment is summarised in the consort diagram (Figure 9).

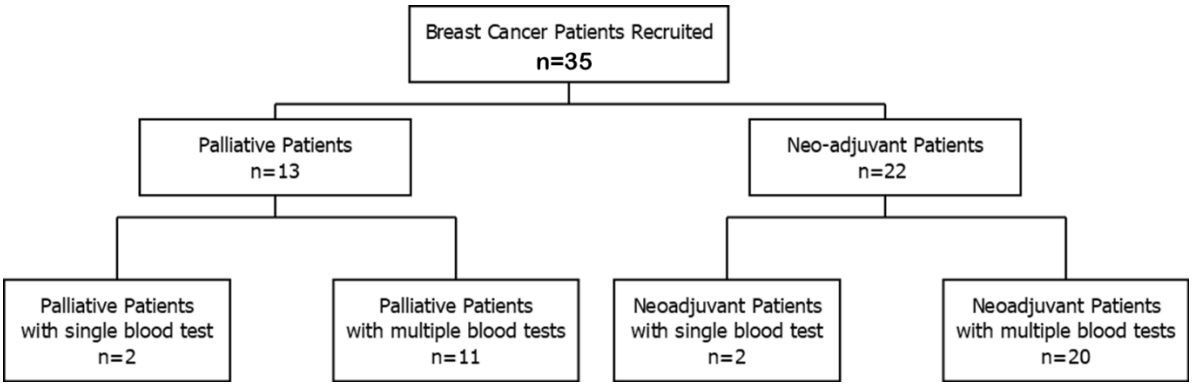


Figure 9 - **Consort Diagram summarising breast cancer patient recruitment cohorts.** Two cohorts are shown, firstly 'palliative patients' with metastatic breast cancer and secondly 'neoadjuvant' patients receiving neoadjuvant chemotherapy with curative intent.

Recruitment of 22 patients undergoing neoadjuvant chemotherapy with curative intent was completed by September 2016. Follow-up bloods including post-operative samples were collected until July 2017. Clinical history was collected until August 2018. 2 patients were lost to follow up or refused follow-up blood sampling. Patient characteristics are summarised in Table 12.

	Palliative Treatment		Neo-adjuvant Treatment	
	No.	%	No.	%
Total	13	100	22	100
Histology				
Ductal	7	54	17	76
Lobular	2	15	1	4
Other / Mixed invasive	4	31	4	20
Lymph node Status				
Positive	9	70	12	55
Negative	2	15	8	35
Unknown	2	15	2	10
Tumour size				
T0,1	3	23	12	55
T2	5	38	4	20
T3	1	8	4	20
T4	0	0	0	0
Unknown	4	31	2	10
Tumour Grade				
I	1	8	2	10
II	2	15	12	55
III	6	46	6	25
Unknown	4	31	2	10
Hormone Receptor				
ER positive	11	84	11	50
ER Negative	2	16	9	40
Unknown	0	0	2	10
HER 2 Status				
Positive	3	23	5	22
Negative	10	77	15	68
Unknown	0	0	2	10

Table 12 – **Summary of Breast Cancer Patient Characteristics**. Palliative staging is for the initial primary tumours. Neoadjuvant tumour size and staging is post neoadjuvant treatment.

3.4 Mutation profiles of matched tissue and blood in palliative patients

3.4.1 Tumour DNA analysis

For each of 13 patients, 20ng of FFPE derived tumour DNA was analysed by sequencing with the AcquiRes ampliseq panel. This is an in-house designed panel covering four genes (*PIK3CA, TP53, ERBB2* and *ESR1)* (Guttery et al., 2015). Seven of the 13 patients (53.8%) had a detected SNV (mutation) with a VAF >5% in tumour DNA. *PIK3CA* was most the most commonly mutated gene, detected in 5 out of 13 patients (38.4%), with 4 of the 13 (30.7%) patients having a *TP53* mutation and 2 patients having a mutation in both genes (Figure 10 and Table 13). Only *PIK3CA* and *TP53* mutations were detected in tumour DNA with this panel.

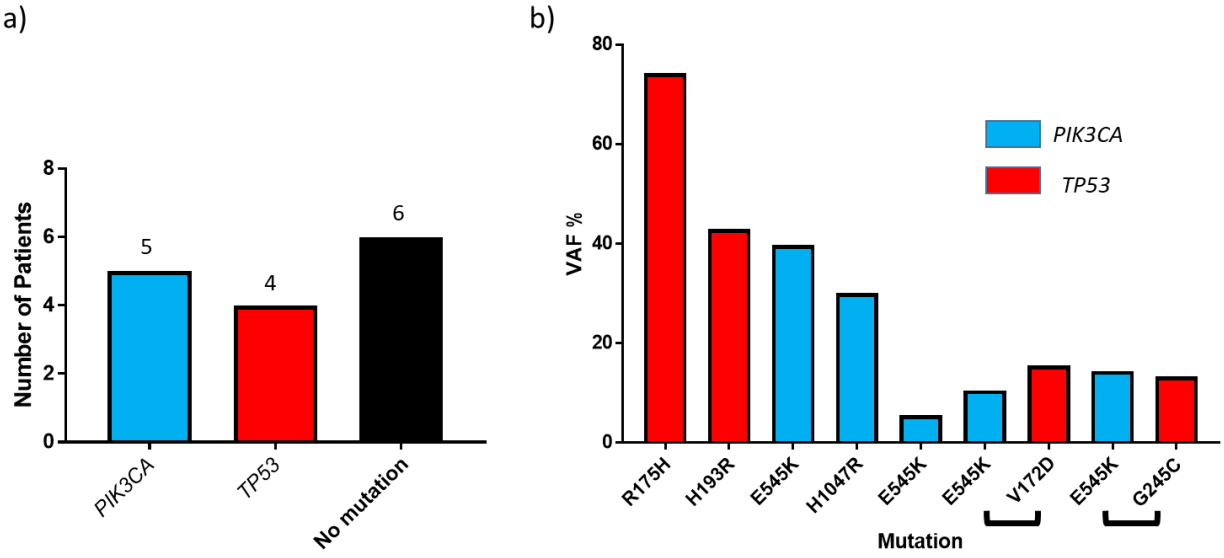


Figure 10 **– Summary of tumour DNA sequencing, palliative cohort**. a) Bar chart showing the cumulative number of patients for each gene that have a mutation. Note some patients have > 1 mutation and are therefore counted twice b) Frequency of specific mutation (> 5% VAF) across all sequenced tumours. Solid bar indicates two detectable mutations in a single tumour tissue.

3.4.2 Total cfDNA analysis – Palliative Patients

In all 7 patients with a mutation detected in tumour DNA, the matched total cfDNA was also analysed by NGS and / or ddPCR. In 5 patients a single variant was analysed and in 2 patients

2 variants were analysed. In total 5 of the 7 patients (71.4%) had the same SNV detected in their blood indicating the presence of ctDNA (Table 13).

	Patient	Tumour Variant Detected	Gene	Tumour DNA variant VAF	Number of Plasma Samples	Total cfDNA positive for tumour specific variant		cfDNA VAF†
						YES/NO	No. +ve Samples	
1	CRB03	YES	*PIK3CA*	31.0% H1047R	1	YES	1	1.0% H1047R
2	CRB08	NO		-	6	-		-
3	CRB09	NO		-	6	-		-
4	CRB17	YES	*PIK3CA* *TP53*	10.4% E545K 15.4% V172D	7	YES	5	0.09-29.4% E545K 0.08-53.1% V172D
5	CRB24	YES	*PIK3CA*	39.7% E545K	10	YES	7	0.7-52.5% E545K
6	CRB25	NO		-	10	-		-
7	CRB27	NO		-	8	-		-
8	CRB37	YES	*PIK3CA*	5.3% E545K	3	NO	0	-
9	CRB68	NO		-	1	-		-
10	CRB72	YES	*PIK3CA*	14.4% E545K	5	NO	0	-
			TP53	13.3% G245C		Not done – no assay	-	-
			PIK3CA	4.5% E524K		Not done VAF <5%	-	-
			TP53	2.8% R213X		Not done VAF <5%	-	-
11	CRB73	YES	*TP53*	42.9% H193R	3	YES	3	2.0-71.4% H193R
12	CRB77	YES	*TP53*	74.0% R175H	8	YES	8	0.05-17.6% R175H
13	CRB91	NO		-	2	-		-
	Total	7			70	5	24	

Table 13 – **Summary of total cfDNA analysis for tumour specific variants.** †For individual cfDNA results see patient specific graphs and Table 15, page 105.

Due to the low amount of cfDNA it was not possible to repeat every sample in duplicate. When ctDNA levels could be detected at high levels by ddPCR the results were highly reproducible. However, at low levels of VAF of ~0.05% results with a one or two positive droplets could be discarded as false negatives. Therefore, all result with only one or two positive droplets were repeat with increased amounts of input DNA to confirm if they are true negative or not. This is shown more clearly in Figure 11 where using increased amount of DNA on replication confirms the negative samples but also finds one previously false negative sample. All samples were confirmed in this way if the amount of cfDNA allowed.

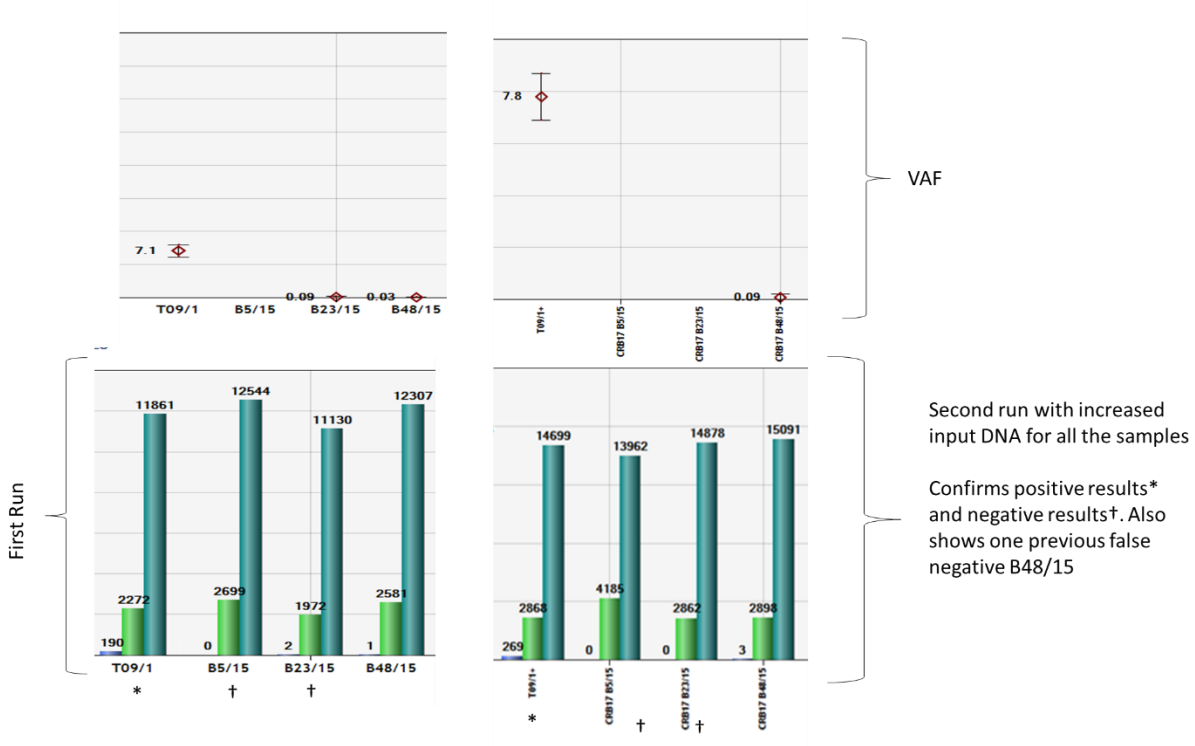


Figure 11 – **Replication of experiments with rare events resulting in low level VAF ~0.05%.** Upper panel; VAF per sample, Lower panel; total number of events (grey bar), positive droplets for mutant *PIK3CA* E545K (blue bar), Wild type droplets for *PIK3CA* E545K (green bar). Left first run, right second run with increased amount of input DNA as shown by increased number of wildtype events (green bar).

3.4.2.1 Patient monitoring by tracking of SNVs in total cfDNA

The 7 patients with a known tumour specific variant >5% VAF (Table 13) are discussed in this section. CRB03 was a special case and is discussed at the end of the section.

Patient CRB17 had two SNV's detected in tumour DNA, *PIK3CA* p.E545K and *TP53* p. V172D (Figure 12). ctDNA was undetectable in the first two blood samples, which were taken after the commencement of capecitabine chemotherapy. Low levels of ctDNA (4 copies/ml) was detected in the third sample (blood 3), taken 14 weeks (104 days) prior to relapse, as evidenced by the detected of both tumour specific SNVs. Blood 4, taken 7 weeks (51 days) prior to a relapse CT scan, showed rising ctDNA (98-135 copies/ml) confirming molecular relapse prior to clinical relapse. Latterly, a second rise in ctDNA levels occurred in the 7th

blood sample 2.5weeks (18 days) before clinical relapse and just prior to death. This shorter lag-time between rise in ctDNA levels and clinical relapse is interesting, as the CT scans from this period suggest there was a sudden increase in the growth rate of the cancer (Figure 13, page 94). A blood test in between bloods 6 and 7 may have predicted when this increase in growth rate occurred. At the time blood test 7 was taken the patient was clinically well. Therefore, it is possible during the 2.5 week window before confirmation of disease progression, further treatment could have been initiated that may have had an impact on clinical outcome.

Patient CRB73 had a *TP53* p.H193R mutation detected in tumour DNA, that was also detected in serial plasma samples (Figure 14, page 95). The second blood sample shows decreasing levels of ctDNA correlating with a partial response on CT, indicating successful treatment with chemotherapy (paclitaxel). Evidence of progressive disease through rising ctDNA level was seen subsequently, 3 weeks (22 days) prior to death. This rise in ctDNA level occurred *during* chemotherapy, suggesting that this therapy was ineffective.

Patient CRB77 had a *TP53* p.R175H mutation detected in tumour DNA, and serial plasma samples (Figure 15, page 96). Bloods 2 and 3 had mutation detection below 1% VAF via NGS, which was also borderline via ddPCR (see Appendix 7.3). ctDNA levels showed a gradual increase during anti-oestrogen therapy, with a more pronounced increase during the use of an ineffective trial drug (R06874281 an anti-FAP/interleukin-2 fusion protein antibody). There was a further marked increase in ctDNA at the final blood sample (blood 7) but no corroborating CT to confirm progression.

Blood number		1	2	3	4	5	6	7
Total cfDNA copies/ml		8030	4545	5000	5000	5606	6384	33181
PIK3CA p. E545K	AF (%)	0	0	0	1.9	13.0	12.3	55.4
	Copies/ml	0	0	0	98	728	785	18382
TP53 p. V172D	AF (%)	0	0	0.08	2.5	24.0	16.3	53.1
	Copies/ml	0	0	4	125	1345	1040	17619

Figure 12 – **Timeline and total cfDNA analysis of patient CRB17.** Top left shows date and pathology results of the initial breast cancer tumour along with the results of tumour DNA NGS. Top centre; timeline of clinical events and timing of blood tests. Centre; Total cfDNA levels in copies/ml (black) and ctDNA as defined by detection of *PIK3CA* p.E545K (blue) and *TP53* p.V172D (red) in copies/ml (red), lagtime (LT) between rise in ctDNA levels and CT scan showing progression shown by solid bars. Bottom; Table showing total cfDNA levels in copies/ml and detected mutation(s) as allele fraction (AF) and copies/ml. SD = Stable disease, PD = progressive disease, PR = partial response.

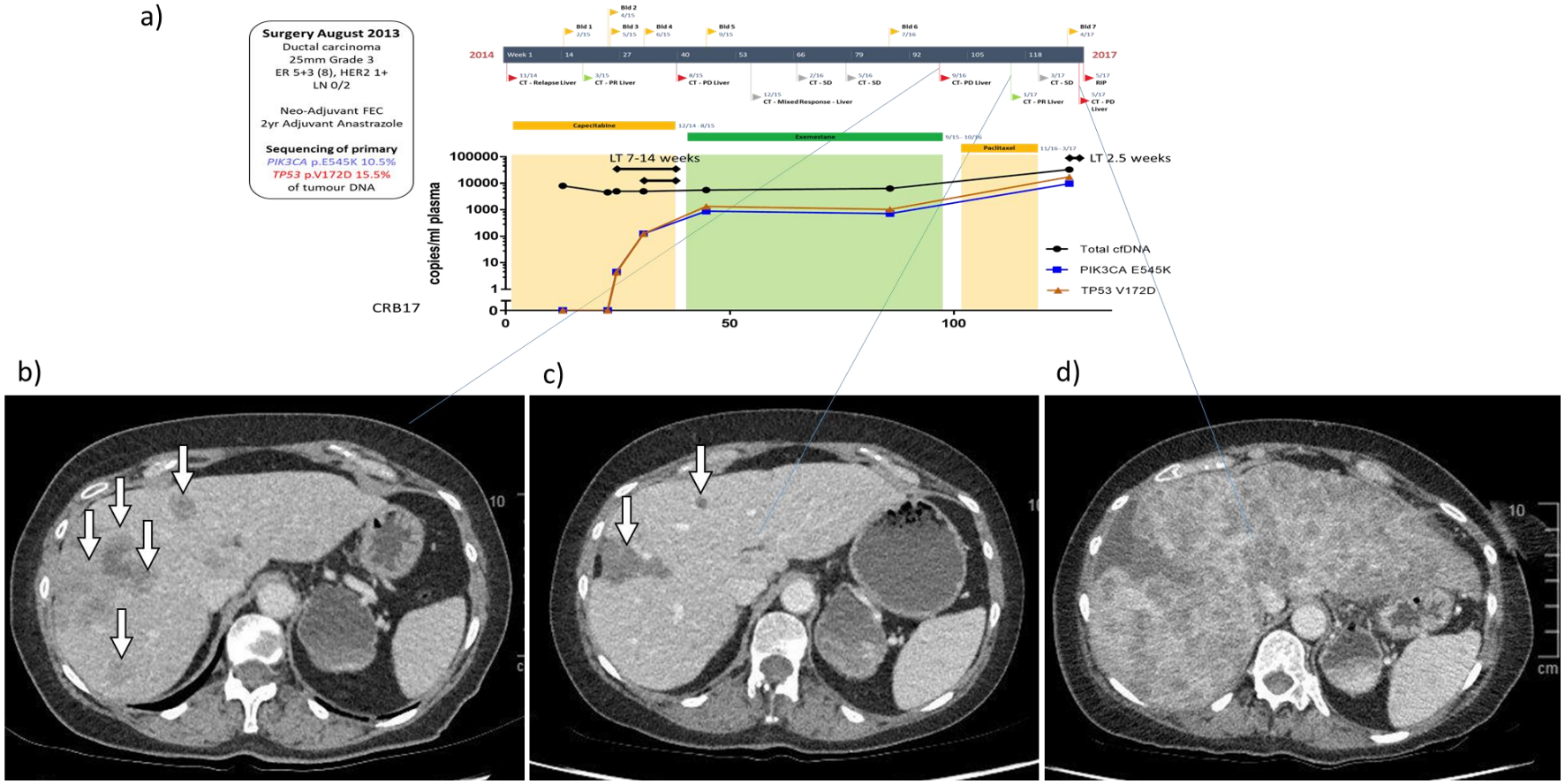
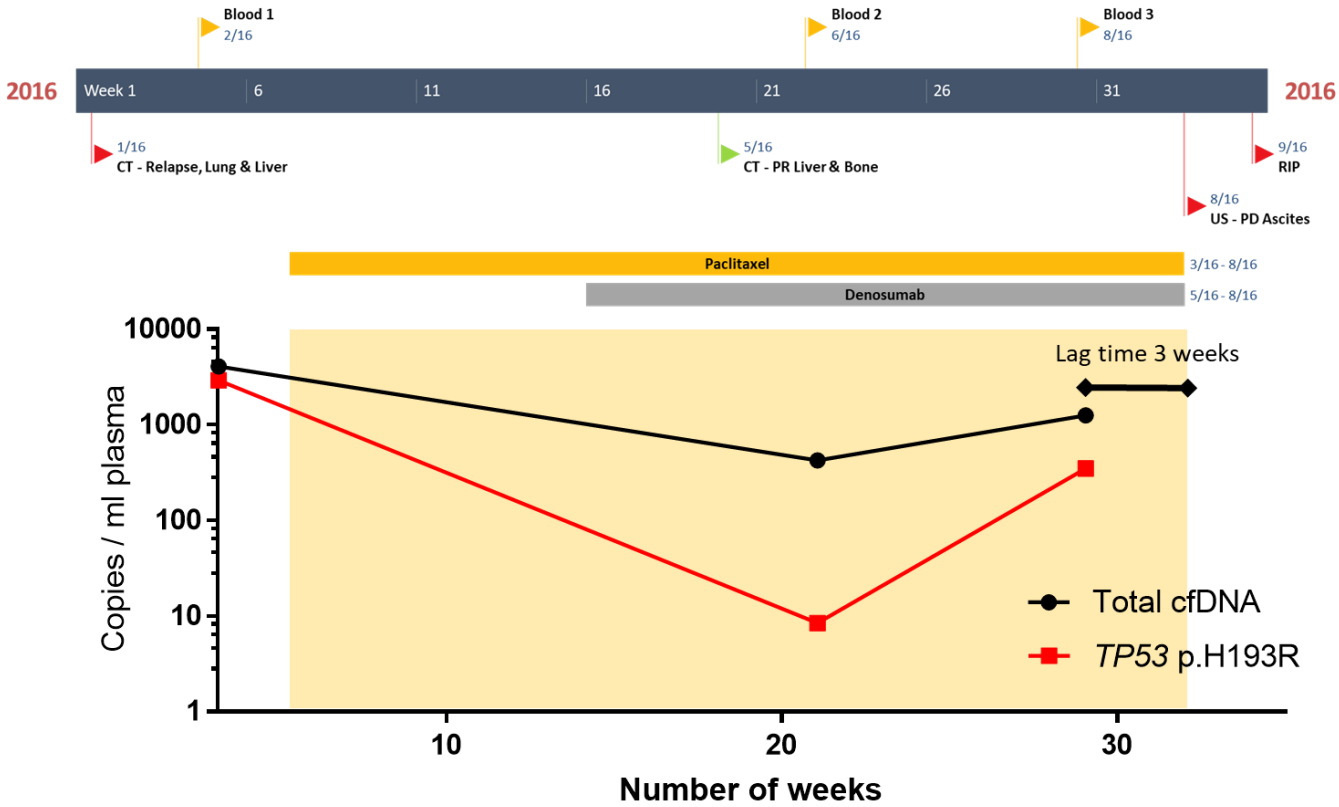


Figure 13 – **Timeline of Patient CRB17 with corresponding CT scans showing disease in the liver.** a) Timeline of patient CRB17 shown in Figure 12. b) CT scan May 2016, showing disease progression of liver metastases, manifesting as multiple diffuse areas of disease within the liver (arrowed). c) CT scan January 2017, showing disease response of metastases within the liver, showing one well-defined triangular area of disease and a further single subcentimetre deposit (arrowed). d) CT scan May 2017, showing progressive disease with extensive diffuse liver involvement

Surgery Sept 2014
Multifocal Ductal carcinoma
Primary 1: 15mm Grade 2
HER2 = 1+, ER = 5+3 (8)
Primary 2: 14mm Grade 3
Her2 = 1+, ER 2+1 (3)
LN 3/22

Neoadjuvant FEC-T
Adjuvant Tamoxifen 2 years

Sequencing of primary
TP53 p.H193R
42.9% of tumour DNA



CRB73

Blood number		1	2	3
Total cfDNA copies/ml		4106	424	1260
TP53 p. H193R	AF (%)	71.4	2.0	27.7
	Copies/ml	2931	8	349

Figure 14 - **Timeline and total cfDNA analysis of patient CRB73.** Top left shows date and pathology results of the initial breast cancer tumour along with the results of tumour DNA NGS. Top centre; timeline of clinical events and timing of blood tests. Centre; total cfDNA levels in copies/ml (black) and ctDNA as defined by detection of *TP53* p.H193R (red) in copies/ml (red), lag time between rise in ctDNA levels and CT scan showing progression shown by solid bars. Bottom; Table showing total cfDNA levels in copies/ml and detected mutation(s) as allele fraction (AF) and copies/ml. PD = progressive disease, PR = partial response.

Blood number		1	2	3	4	5	6	7
Total cfDNA copies/ml		909	606	909	1515	696	975	821
TP53 p. R175H	AF (%)	8.3	0.05	0.07	0.05	0.27	1.4	17.6
	Copies/ml	75.5	0.3	0.6	0.7	1.8	13.7	147

Figure 15- **Timeline and total cfDNA analysis of patient CRB77.** Top left shows date and pathology results of the initial breast cancer tumour along with the results of tumour DNA NGS. Top centre; timeline of clinical events and timing of blood tests. Centre; Total cfDNA levels in copies/ml (black) and ctDNA as defined by detection of *TP53* p.R175H (red) in copies/ml (red). Bottom; Table showing total cfDNA levels in copies/ml and detected mutation(s) as allele fraction (AF) and copies/ml. SD = stable disease, PD = progressive disease, PR = partial response, LN = lymph node.

For patient CRB24 10 blood samples were available for analysis from diagnosis of disease relapse to death. The tumour DNA had a *PIK3CA* p.E454K mutation, which also was detected in 7 of the 10 blood samples (Figure 16). ctDNA was detected in the first sample (1.4 % VAF, 57 copies/ml) but became undetectable during primary therapy for metastatic bone disease suggesting disease response. However, the first CT scan showed stable disease, potentially because CT scans are poor predictors of bone disease. By the time of blood 4, ctDNA levels were exceptionally high (29,829 copies / ml). A CT scan was completed 3 weeks (23 days) later, October 2015, and was reported as disease progression but the oncologists involved in the case were not sufficiently convinced that the current drug therapy was ineffective. At this point the patient was clinically well and the lung changes were consistent with infection as well as metastatic lung disease infiltrating the lymphatic system of the lungs (lymphangitic carcinomatosis) (Figure 17b). The oncologists therefore continued the therapy for another 7 weeks (45 days) before a further CT scan showed definitive disease progression within the lung and appearance of a right sided pleural effusion (Figure 16 & Figure 17c). At this point the patient has become clinically breathless due to this disease. In this case the knowledge of ctDNA would have been useful to the treating doctors. Had the rising ctDNA levels been known to the doctor when blood 4 was taken, this could have indicated the CT findings were more likely to be disease related rather than due to infection. Therefore, the therapeutic regimen (trastuzumab and pertuzumab) may have been revised 10 weeks (68 days) prior to the actual stop date (see Figure 16).

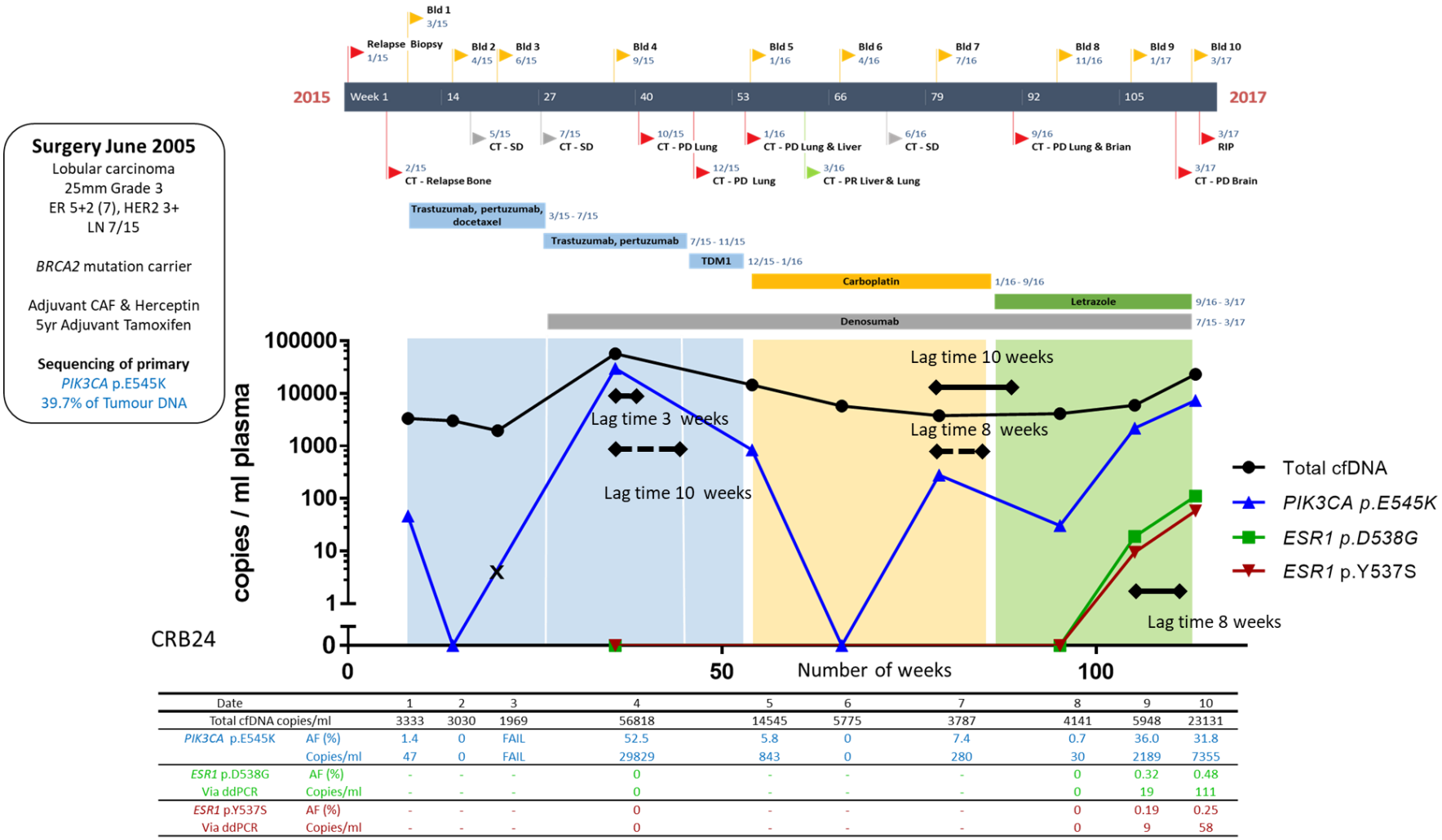
Following confirmation of disease progression on CT (December 2015), the patient was given a course of TDM1, which was clinically ineffective with the subsequent CT (January 2016) showing progressive disease (Figure 17d). There was a decrease in levels of ctDNA during this period (blood 5) but overall they remained elevated (Figure 16). The reason for this is unknown

Following a switch to carboplatin therapy after blood 5 was taken a significant response of the lung disease was seen on CT (March 2016). There was a correlating drop in ctDNA to undetectable levels (blood 6, Figure 16, Figure 17e).

Detectable ctDNA re-emerged in the 7th blood sample, 8 weeks (56 days) prior to a change in therapy to letrazole and 10 weeks (72 days) before a CT scan (September 2016) confirmed disease progression. The

letrazole therapy led to a drop in ctDNA levels (blood 8) but there was no corroborating CT scan. The ctDNA levels rose on the 9th blood sample, 8 weeks (57 days), before disease progression was confirmed. Lastly at the same time (blood 9) there was emergence of two *ESR1* mutations (p.D538G and p.Y537S). This shows emergence of resistance to anti-oestrogen therapy via polyclonal *ESR1* mutations.

CRB72 had two detectable mutations in tumour DNA. A *PIK3CA* p.E545K mutation and a *TP53* p.G245C mutation The bloods were analysed via ddPCR for the *PIK3CA* mutation all of which were negative (Figure 22, page 108). Unfortunately, a ddPCR assay was not available for the G245C mutation. CRB37 also had a detectable *PIK3CA* p.E545K mutation in the tumour DNA and no evidence of the mutation in the three blood samples (Figure 18).



Date		1	2	3	4	5	6	7	8	9	10
Total cfDNA copies/ml		3333	3030	1969	56818	14545	5775	3787	4141	5948	23131
PIK3CA p.E545K	AF (%)	1.4	0	FAIL	52.5	5.8	0	7.4	0.7	36.0	31.8
	Copies/ml	47	0	FAIL	29829	843	0	280	30	2189	7355
ESR1 p.D538G	AF (%)	-	-	-	0	-	-	-	0	0.32	0.48
Via ddPCR	Copies/ml	-	-	-	0	-	-	-	0	19	111
ESR1 p.Y537S	AF (%)	-	-	-	0	-	-	-	0	0.19	0.25
Via ddPCR	Copies/ml	-	-	-	0	-	-	-	0	9	58

Figure 16 - **Timeline and total cfDNA analysis of patient CRB24.** Top left shows date and pathology results of the initial breast cancer tumour along with the results of tumour DNA NGS. Top centre; timeline of clinical events and timing of blood tests. Centre; Total cfDNA levels in copies/ml (black) and ctDNA as defined by detection of *PIK3CA* p.E454K (blue). Lag time between rise in ctDNA levels and CT scan showing progression shown by solid bars, lag time between rise in ctDNA and change in therapy (where different) shown as dashed bar. Emergent *ESR1* p.D538G mutation (green) and p.Y537S (dark red). Bottom; Table showing total cfDNA levels in copies/ml and detected mutation(s) as allele fraction (AF) and copies/ml. SD = stable disease, PD = progressive disease, PR = partial response.

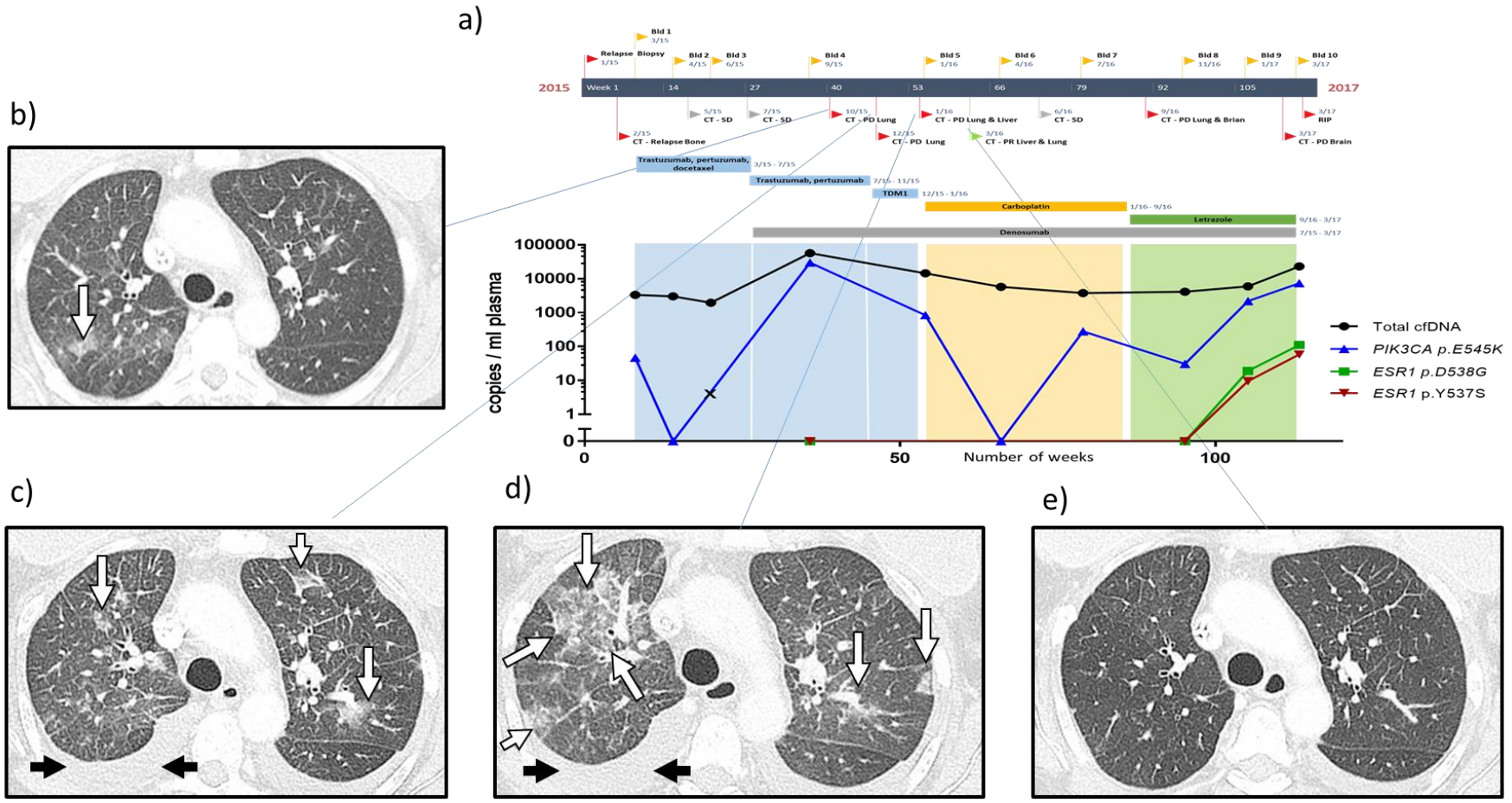
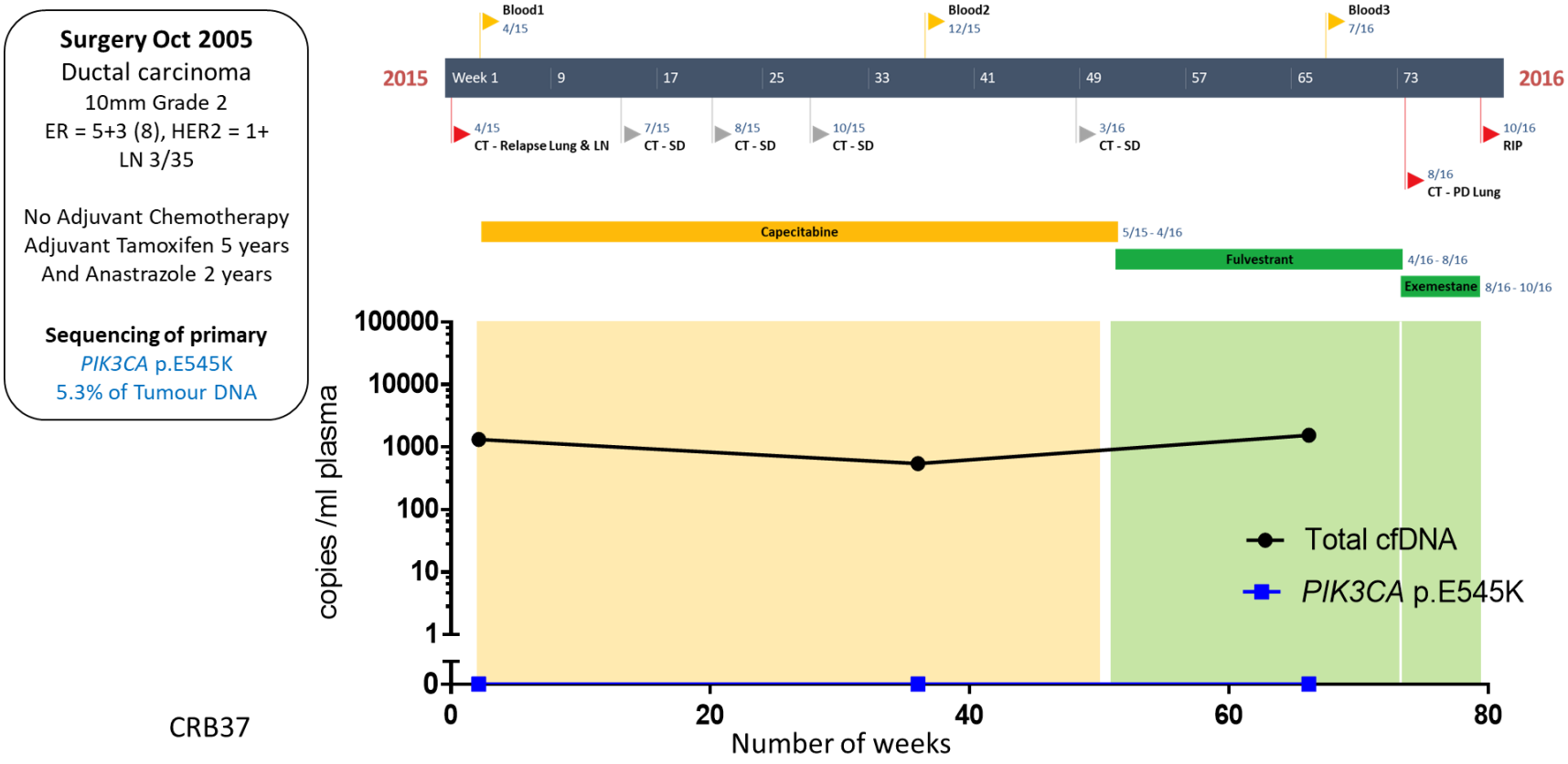


Figure 17 – **Timeline of Patient CRB24 with corresponding CT scans showing lung lymphangitic carcinomatosis.** a) Timeline of patient CRB24. b) CT scan October 2015, showing disease progression that is potentially lymphangitic carcinomatosis or an area of infection (arrowed). c) CT scan December 2015, showing areas of lymphangitic carcinomatosis within the lung (arrowed white) and pleural effusion (arrowed black). d) CT scan January 2016, showing progressive disease with extensive lung lymphangitic carcinomatosis (arrowed white) and pleural effusion (arrowed black). e) CT scan March 2016, showing resolution of lymphangitic carcinomatosis, and pleural effusion following carboplatin therapy.



Blood number		1	2	3
Total cfDNA copies/ml		1333	545	1560
PIK3CA p.E545K	AF (%)	0	0	0
	Copies/ml	0	0	0

Figure 18 - **Timeline and total cfDNA analysis of patient CRB37.** Top left shows date and pathology results of the initial breast cancer tumour along with the results of tumour DNA NGS. Top centre; timeline of clinical events and timing of blood tests. Centre; Total cfDNA levels in copies/ml (black) and ctDNA as defined by *PIK3CA* p.E545K (blue) not detected. Bottom; Table showing total cfDNA levels in copies/ml and detected mutation(s) as allele fraction (AF) and copies/ml. SD = stable disease, PD = progressive disease, LN = lymph node.

Patient CRB03 had a *PIK3CA* p.H1047R mutation in tumour DNA. This patient had a single blood sample. However other tissue and ascitic fluid samples were taken. DNA from all of these samples was analysed by NGS comprising; two tumour regions and a lymph node from the curative resection taken in 1998, two regions from a pleural biopsy taken in 2011 at relapse, and two ascitic samples available from October and November 2014 prior to the patient's death. The *PIK3CA* p.H1047R mutation was detected in all samples (Figure 19). Of note there was a high mutation fraction in the ascitic sample just prior to the patient's death.

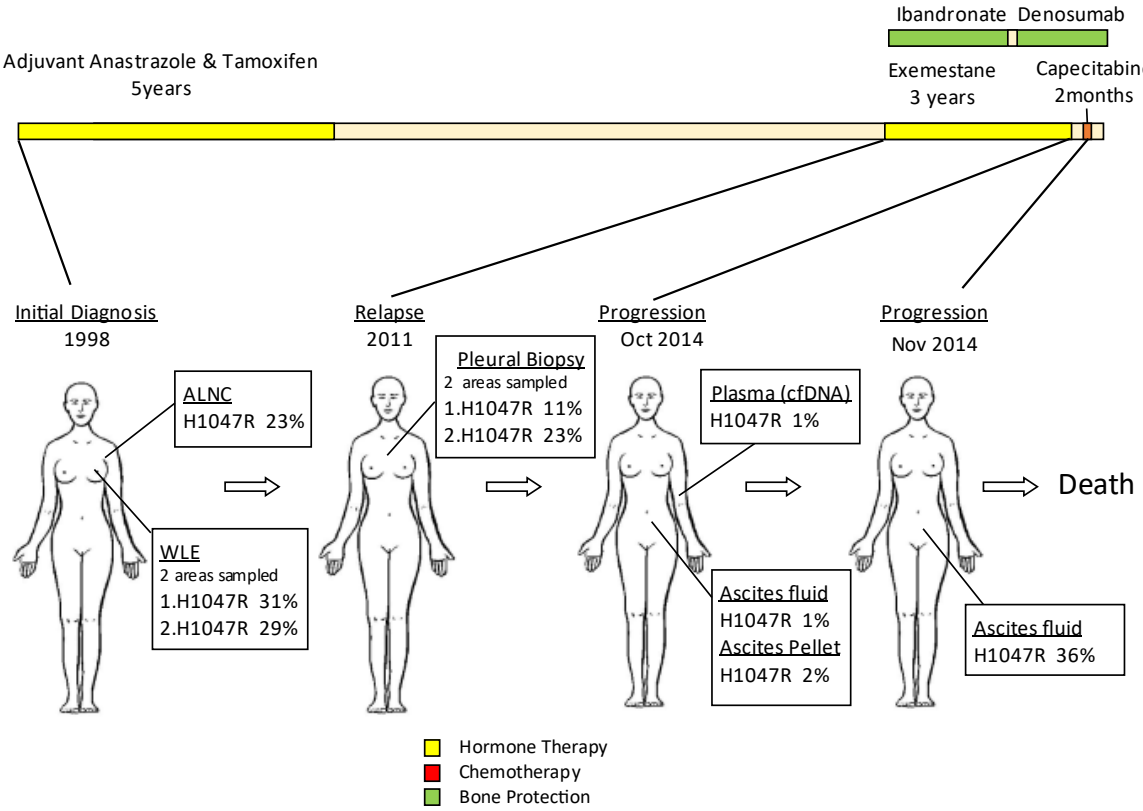


Figure 19 – **NGS results and treatment timeline of CRB03**. Percentages given are VAF.

3.4.3 ESR1 mutation analysis

Based on the results from plasma profiling, specifically CRB24 which showed emergence of two *ESR1* mutations, an additional test was used to screen for the emergence of two common resistance *ESR1* mutations (p. Y537S and p.D538G). This was carried out via ddPCR using the final blood sample collected from each patient. This identified a further 4 patients

that were ctDNA positive, including 3 patients who did not have any tumour SNVs detected at >5% VAF in tumour DNA (Table 14). Subsequently, the other blood samples from these 4 patients were also analysed for the same mutation. For complete cfDNA analysis results and corresponding VAFs see Table 15, page 105.

	CRB	Tumour Variant Detected	Number of Plasma Samples	Total cfDNA positive for tumour specific variant		Total cfDNA positive for ESR1 Mutation		ctDNA Detected at any timepoint?
				YES/NO	No. Samples	YES/NO	No. Samples	YES/NO
1	CRB003	YES	1	YES	1	NO	0	YES
2	CRB008	NO	6		-	YES	2	YES
3	CRB009	NO	6		-	NO	0	NO
4	CRB017	YES	7	YES	5	NO	0	YES
5	CRB024	YES	10	YES	7	YES	2	YES
6	CRB025	NO	10		-	YES	9	YES
7	CRB027	NO	8		-	NO	0	NO
8	CRB037	YES	3	NO	0	NO	0	NO
9	CRB068	NO	1		-	NO	0	NO
10	CRB072	YES	5	NO	0	YES	1	YES
11	CRB073	YES	3	YES	3	NO	0	YES
12	CRB077	YES	8	YES	8	NO	0	YES
13	CRB091	NO	2		-	YES	1	YES
	Total	7	70	5	24	5	15	9

Table 14 – **Summary of tumour DNA NGS, and total cfDNA analysis for tumour specific variant and *ESR1* mutations.**

In the case of CRB08 an emergent *ESR1* p. D538G mutation was detected. This mutation was present during exemestane therapy and likely arose during prior letrazole therapy (Figure 20, page 106). Studies show that these *ESR1* mutations result in a constitutively activated ER receptor and therefore predict failure of aromatase inhibitors (AIs) such as letrazole and exemestane (Jeselsohn et al., 2014). However, it is interesting to note that a change in the AI therapy from letrazole to exemestane elicited a response in this patient despite the presence of an *ESR1* mutation. A similar pattern of emergence of an *ESR1* p.D538G mutation during letrazole therapy, predicating resistance to AI was seen from analysis of patient CRB72 (Figure 22, page 108)

Both *ESR1* mutations were detected in the plasma of patient CRB25. Analysis of plasma from the other available time points showed that this patient had significant levels of ctDNA most bloods (Figure 21). Although the *ESR1* mutations were present in the first sample taken, they were not detected by NGS in the original tumour DNA. However, the patient had previously received adjuvant anastrazole and palliative tamoxifen, either of which may have selected for mutation emergence. As well as predicting resistance to AI therapy, the levels of the two *ESR1* mutations fluctuated with response to chemotherapy and rose when disease became resistant to therapy. Most interestingly, during fulvestrant therapy, both mutations increased, but the *ESR1* p.Y537S mutation was more prevalent than the p.D538G one. This suggested first, that fulvestrant therapy was ineffective for this patient, and secondly that a differential response of the two mutations had occurred meaning that there were potentially different clones of tumour cells containing different mutations.

Rising levels of the two mutations were seen 10 weeks before progression of disease following both blood 3 (lag time 69 days) and blood 9 (67 days). A good response was also seen in response to eribulin therapy, with undetectable ctDNA levels following a single cycle of therapy. Unfortunately, eribulin could not be continued due to drug toxicity, but if the response had been known, a further cycle at a reduced dose may have been indicated. Finally, non-response to epirubicin and carboplatin therapy was seen on ctDNA analysis just before death (blood 10), despite a CT-scan showing a partial response. This indicates that ctDNA detection may provide additional information on disease response, in certain circumstances compared to CT.

Patient CRB91, had a low level *ESR1* p.D538G mutation (0.28% VAF, 7 copies/ml) detected on diagnosis of metastatic disease (Figure 23). During treatment with chemotherapy, a second blood sample was taken and the mutation was undetectable. The patient later relapsed with metastatic brain disease before death. There was no warning of this from this ctDNA analysis.

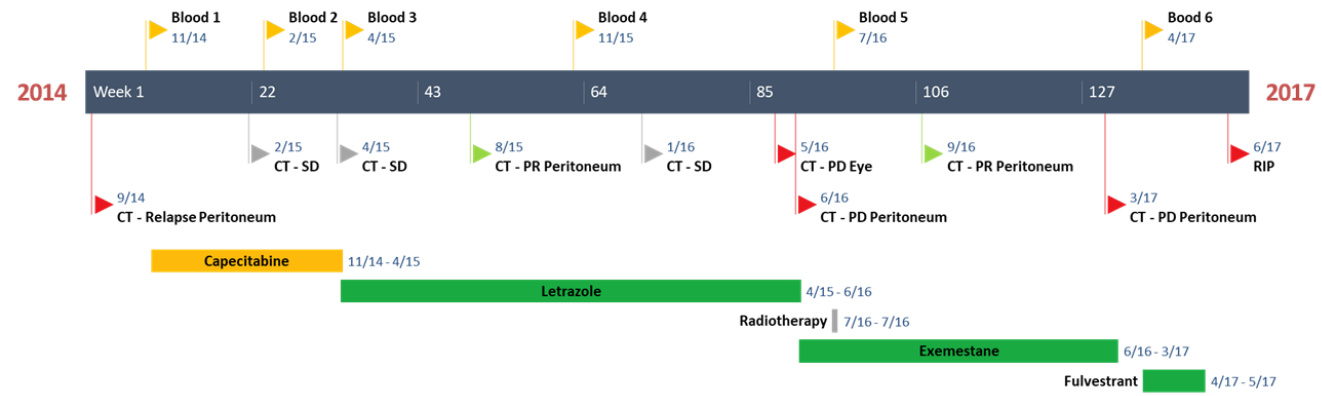
		Tumour Variant			cfDNA Variant		Blood number and gene specific VAFs									
	Patient	Gene	Mutation	VAF	Gene	Mutation	1	2	3	4	5	6	7	8	9	10
1	CRB003	*PIK3CA*	H1047R	31.0%	*PIK3CA*	H1047R	1.0	-	-	-	-	-	-	-	-	-
					ESR1	D538G	0	-	-	-	-	-	-	-	-	-
					ESR1	Y537S	0	-	-	-	-	-	-	-	-	-
2	CRB008	No Mutation			*ESR1*	D538G	FAIL	0	0	0	0.63	0.44	-	-	-	-
					ESR1	Y537S	*	*	*	*	*	0	-	-	-	-
3	CRB009	No Mutation			*ESR1*	D538G	*	*	*	*	*	0	-	-	-	-
					ESR1	Y537S	*	*	*	*	*	0	-	-	-	-
4	CRB017	*PIK3CA*	E545K	10.4%	*PIK3CA*	E545K	0	0	0.09	2.5	15.9	11.3	29.4	-	-	-
		TP53	V172D	15.4%	*TP53*	V172D	0	0	0.09	2.5	24.0	16.3	53.1	-	-	-
					ESR1	D538G	*	*	*	*	*	*	0	-	-	-
					ESR1	Y537S	*	*	*	*	*	*	0	-	-	-
5	CRB024	*PIK3CA*	E545K	39.7%	*PIK3CA*	E545K	1.4	0	FAIL	52.5	5.8	0	7.4	0.7	36.0	31.8
					ESR1	D538G	*	*	*	0	*	*	*	0	0.32	0.48
					ESR1	Y537S	*	*	*	0	*	*	*	0	0.19	0.25
6	CRB025	No Mutation			*ESR1*	D538G	0.64	0	0.53	3.3	0	0	0.37	0.24	0.59	0.48
					ESR1	Y537S	6.0	0	2.1	10.5	0.07	0.13	10.2	1.1	3.6	9.5
7	CRB027	No Mutation			*ESR1*	D538G	*	*	*	*	*	*	*	0	-	-
					ESR1	Y537S	*	*	*	*	*	*	*	0	-	-
8	CRB037	*PIK3CA*	E545K	5.3%	*PIK3CA*	E545K	0	0	0	-	-	-	-	-	-	-
					ESR1	D538G	*	*	0	-	-	-	-	-	-	-
					ESR1	Y537S	*	*	0	-	-	-	-	-	-	-
9	CRB068	No Mutation			*ESR1*	D538G	0	-	-	-	-	-	-	-	-	-
					ESR1	Y537S	0	-	-	-	-	-	-	-	-	-
10	CRB072	*PIK3CA*	E545K	14.4%	*PIK3CA*	E545K	0	0	0	0	0	-	-	-	-	-
		TP53	G245C	13.3%	*TP53*	G245C	Not tested – no ddPCR assay									
		PIK3CA	E524K	4.5%	*PIK3CA*	E524K	Below VAF threshold									
		TP53	R213X	2.8%	*TP53*	R213X	Below VAF threshold									
					ESR1	D538G	0	0	0	0	0.23	-	-	-	-	-
					ESR1	Y537S	*	*	*	*	0	-	-	-	-	-
11	CRB073	*TP53*	H193R	42.9%	*TP53*	H193R	71.4	2.0	27.7	-	-	-	-	-	-	-
					ESR1	D538G	*	*	0							
					ESR1	Y537S	*	*	0							
12	CRB077	*TP53*	R175H	74.0%	*TP53*	R175H	8.3	0.05	0.07	0.05	0.27	1.4	17.6	-	-	-
					ESR1	D538G	*	*	*	*	*	*	0			
					ESR1	Y537S	*	*	*	*	*	*	0			
13	CRB091	No Mutation			*ESR1*	D538G	0.28	0	-	-	-	-	-	-	-	-
					ESR1	Y537S	0	*	-	-	-	-	-	-	-	-

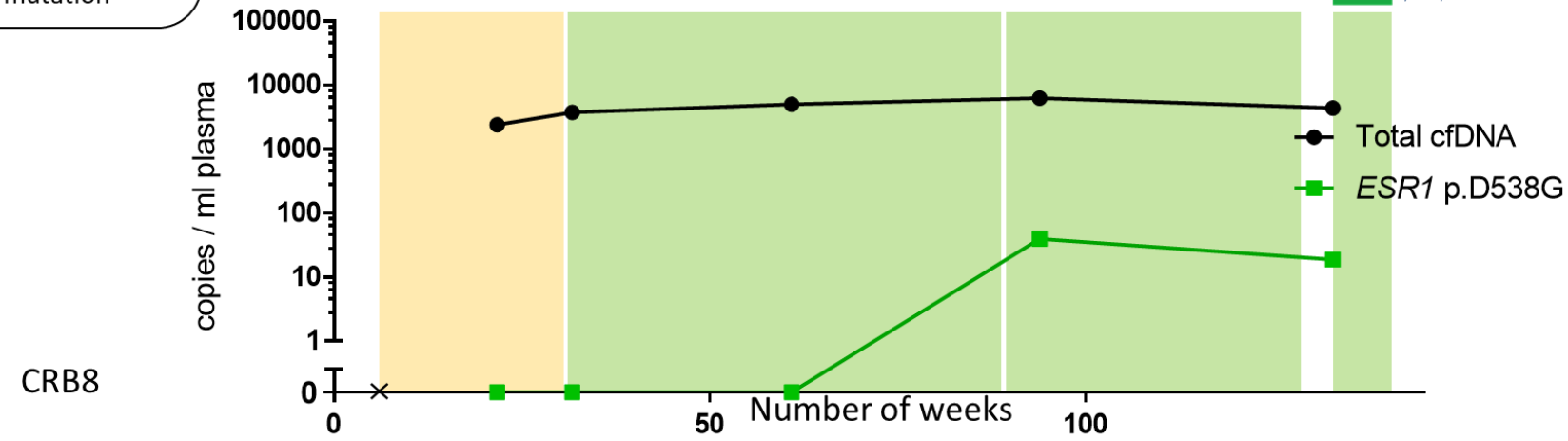
Table 15 – **Palliative breast patients: Complete results of all total cfDNA NGS and ddPCR analyses**. (-) no blood sample (*) sample not tested.

Surgery May 2005
Lobular carcinoma
40mm Grade 3
ER +ve (strong) HER2 0
LN 0/17

Neo-Adjuvant CAF-T
5yr Adjuvant Anastrazole

Sequencing of primary
No mutation



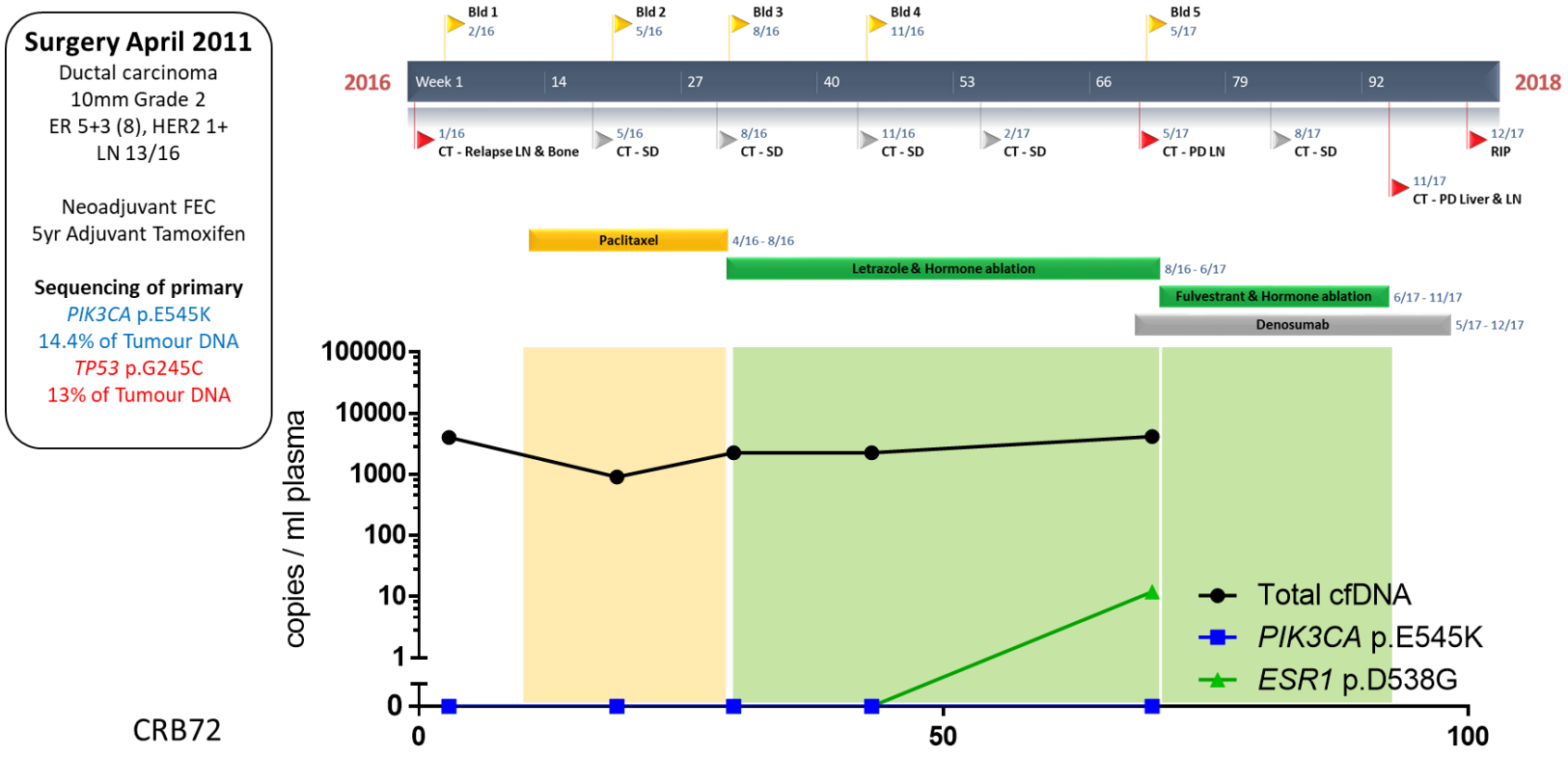


Blood		1	2	3	4	5	6
Total cfDNA copies/ml of plasma		Fail	2424	3787	5045	6333	4439
ESR1 p. D538G	%	-	0	0	0	0.63	0.44
	Copies/ml	-	0	0	0	40	19

Figure 20 – **Timeline and total cfDNA analysis of patient CRB08.** Top left shows date and pathology results and tumour DNA NGS of the initial breast cancer. Top centre; timeline of clinical events and blood tests. Centre; Total cfDNA levels in copies/ml (black) and ctDNA as defined by detection of *ESR1* p.D538G (green) in copies/ml. Bottom; Table showing total cfDNA levels in copies/ml and detected mutation(s) as allele fraction (AF) and copies/ml. SD = Stable disease, PD = progressive disease, PR = partial response.

Blood number		1	2	3	4	5	6	7	8	9	10
Total cfDNA copies/ml		2020	1287	5455	11363	3469	12363	28166	6643	7477	23560
ESR1 p.D538G	AF (%)	0.64	0	0.53	3.3	0	0	0.37	0.24	0.59	0.48
	Copies/ml	6	0	29	375	0	0	104	16	44	113
ESR1 p.Y537S	AF (%)	6	0	2.1	10.5	0.07	0.13	10.2	1.1	3.6	9.5
	Copies/ml	121	0	114	1193	2	16	2873	73	269	2238

Figure 21 – **Timeline and total cfDNA analysis of patient CRB25**. Top left shows date and pathology results and tumour DNA NGS of the initial breast cancer tumour . Top centre; timeline of clinical events and blood tests. Centre; Total cfDNA levels in copies/ml (black) and ctDNA as defined by detection of *ESR1* p.D538G (green) and p.Y537S (dark red), lagtime between rise in ctDNA levels and CT scan showing progression shown by solid bars. Bottom; Table showing total cfDNA levels in copies/ml and detected mutation(s) as allele fraction (AF) and copies/ml. SD = Stable disease, PD = progressive disease, PR = partial response.



Blood number		1	2	3	4	5
Total cfDNA copies/ml		4060	909	2272	2284	4181
PIK3CA p. E545K	AF (%)	0	0	0	0	0
	Copies/ml	0	0	0	0	0
ESR1 p.D538G	AF (%)	0	0	0	0	0.23
	Copies/ml	0	0	0	0	13

Figure 22 - **Timeline and total cfDNA analysis of patient CRB72.** Top left shows date and pathology results and tumour DNA NGS of the initial breast cancer. Top centre; timeline of clinical events and blood tests. Centre; Total cfDNA levels in copies/ml (black) and ctDNA ; *PIK3CA* p.E545K (blue), *ESR1* p.D538G (green). The *TP53* p.G245C mutation was not analysed in total cfDNA. Bottom; Table showing total cfDNA levels in copies/ml and detected mutation(s) as allele fraction (AF) and copies/ml. SD = stable disease, PD = progressive disease, LN = lymph node.

Blood		1	2
Total cfDNA copies/ml of plasma		2626	1313
ESR1 p. D538G	%	0.28	0
	Copies/ml	7	0

Figure 23– **Timeline and total cfDNA analysis of patient CRB25**. Top left shows date and pathology results of the initial breast cancer tumour along with the results of AcquiRes sequencing of tumour DNA. Top centre; timeline of clinical events and timing of total cfDNA samples from blood tests. Centre; Total cfDNA levels in copies/ml (black) and ctDNA as defined by detection of *ESR1* p.D538G (green). Bottom; Table showing actual DNA levels - total cfDNA levels in copies/ml and detected mutation(s) as allele fraction (AF) and copies/ml. SD = Stable disease, PD = progressive disease, PR = partial response, LN = lymphnode.

3.4.4 Results: Neoadjuvant patients

During the project a new IonTorrent technology (Oncomine™), developed for profiling cfDNA became available. This approach involves individual barcoding of DNA molecules before amplification PCR, resulting in increased sensitivity of detection of variants down to 0.1% from 10ng template DNA. Unfortunately, it was not possible to analyse all patient samples using this technology due to the high cost (>£350 per sample). Therefore, it was elected that those patients who had two blood tests, taken before and after their neoadjuvant chemotherapy would be analysed, to monitor the effects of this treatment. 10 patients fitted the criteria (Figure 24). This was carried out in the lab by Dr K Page, a senior research fellow in the Shaw lab (Table 16).

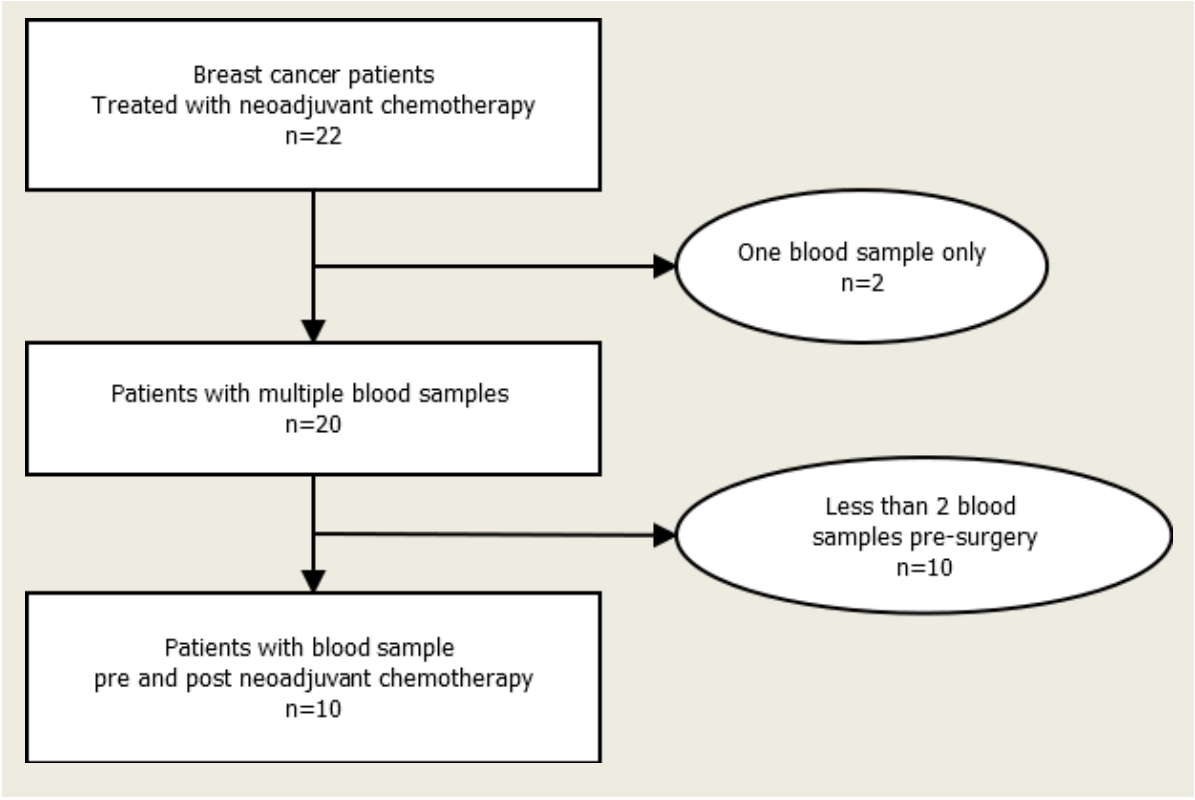


Figure 24 – **Consort diagram showing workflow analysis of patients treated with neoadjuvant chemotherapy.** Only those patients with two or bloods pre-surgery were analysed. Where more than two bloods were available the pre-chemotherapy and pre-surgical blood was analysed.

Patient	Bld	Relapse?	No. reads	Mean Depth	Mutation	Gene	Freq (VAF%	Quality	Threshold* (VAF%)	No. Reads	Mol. Fam.
CRB001	1	NO	72,991	2,247	none						
	2		62,087	2,336	none						
CRB020	1	NO	75,968	2,339	none						
	2		89,388	3,112	p.E545K	*PIK3CA*	0.09	10.3	0.05	3348	3
CRB032	1	NO	78,530	2,971	none						
	2		74,763	2,908	none						
CRB050	1	NO	26,388	774	none						
	2		66,386	3,624	p.R282W	*TP53*	0.07	7.1	0.1	2990	2
CRB053	1	YES	83,337	2,687	p.R213Q	*TP53*	0.07	7.2	0.1	2987	2
					p.C141R	*TP53*	0.16	22	0.05	3168	5
	2		60,131	1,937	p.C141R	*TP53*	0.18	22	0.1	2175	4
CRB092	1	NO	72,726	1,836	none						
	2		65,045	2,586	none						
CRB097	1	YES	97,638	2,223	p.E545K	*PIK3CA*	1.44	252.3	0.1	1668	24
	2		86,448	1,960	none						
CRB098	1	NO	91,762	1,608	none						
	2		108,258	2,574	none						
CRB100	1	NO	98,450	2,771	none						
	2		91,520	1,986	none						
CRB104	1	NO	71,747	1,716	p.I195T	*TP53*	0.12	12.2	0.1	1697	2
	2		81,143	2,799	none						

Table 16 – **Summary of Results from Oncomine Breast Cancer panel version 2** - Analysis of Neoadjuvant patients. Bloods containing detected variants in sequenced genes are shown as white cells on the right of the table. *minimum VAF threshold above which a variant is called, a value that alters depending on read depth. Mol. Fam. = molecular families.

Using the manufacturers settings for detection of ctDNA; VAF > calculated threshold in 2+ molecular families, 5 of the 10 patients (50%) had a detectable mutation in ctDNA. In 2 of these patients (CRB97 and 104) the mutation resolved following neoadjuvant chemotherapy, while in 3 patients the variants either persisted or became apparent before surgery. In the case of CRB53 the *TP53* p.C141R mutation persisted, whereas the low frequency p.R213Q mutation became undetectable. In the case of CRB20 and CRB50 a mutation became apparent following neoadjuvant chemotherapy.

Of interest, 2 of the 5 patients (40%) who had detectable ctDNA (CRB53 + 97) have relapsed whereas none (0%) of the patients without detected ctDNA have relapsed. These 2 patients

had the two highest VAFs in ctDNA. Patient CRB53 relapsed with metastatic liver disease 67 weeks (468 days) after surgery with curative intent. Patient CRB97 relapsed with brain disease 31 weeks (220 days) after curative surgery. The time between the first ctDNA positive blood sample and relapse was 93 weeks (656 days) and 52 weeks (368 days) respectively.

3.4.5 Discussion: Breast cancer and ctDNA

3.4.5.1 ctDNA in patients with metastatic breast cancer: Palliative cohort

Using a combination of panel based NGS of FFPE primary tumour DNA, and personalised mutation tracking by ddPCR and NGS, ctDNA was detected in 9 of 13 patients with metastatic breast cancer (69.2%) who were followed with serial blood sampling (Table 15, page 105). This frequency of detection is consistent with other studies using "off the shelf" NGS mutation panels (Fribbens et al., 2018, Page et al., 2017). Nine patients had ctDNA mutations in *TP53, PIK3CA,* and *ESR1*. In all patients where ctDNA was detected, ctDNA levels tracked with treatment response, showing increase at time of disease progression, and decrease or resolution at times of disease response. In the majority of cases, rising ctDNA preceded disease relapse, by up to 10 weeks. This potentially provides a window of opportunity for switching anti-cancer therapy prior to clinical disease progression on CT. For example, patients CRB17 and CRB73 had additional avenues of treatment still available to them at the time of progression (Cardoso et al., 2018) just before death. Therefore, there is a potential to identify a time period between rise in ctDNA levels before radiological detection of disease progression occurs, during which there may be a clinical opportunity to instigate a different anti-cancer therapy. This information would be valuable to treating clinicians if it could be provided in a timely manner to the clinic and highlights the importance of investigating, via clinical trials, whether acting on such information improves patient overall survival as is being carried out in plasmaMATCH (CRUK, 2018b) and other trials.

The frequency of blood taken in this study was every 3-6 months. In a number of patients, it is apparent that, had a blood sample been taken in the 10-week window before progression or following treatment, changes in ctDNA levels may have predicted progression or response to a greater degree than was possible in this study. This is an argument for more regular blood sampling during patient monitoring to help detect disease progression.

The analysis of 2 common *ESR1* gene mutations, identified emergent mutations in 5 of 11 patients (45.5%). The emergence of these mutations is linked with resistance to anti-oestrogen therapy, particularly AIs, showing that ctDNA can be used to detect the mechanism of resistance to therapy as well as predicting tumour growth and response. However, response to a switch in AI therapy was seen in patient CRB08 despite presence of an *ESR1* mutation suggesting that presence of an *ESR1* mutation in total cfDNA may not predict resistance to all AIs equally. This may be related to their slightly different drug properties (Miller et al., 2008). There is also evidence that response to fulvestrant may be somewhat improved in patients with an *ESR1* mutation compared to AI therapy (Fribbens et al., 2016), although patient CRB25 (Figure 21, page 107) showed no response to fulvestrant in the monitoring period. Fulvestrant is an anti-oestrogen therapy which works by binding the receptor and preventing dimerization (Carlson, 2005). The lack of response to fulvestrant in the case of CRB25 suggests that when multiple lines of chemotherapy are still available and a patient has an emergent *ESR1* mutation, chemotherapeutic agents may be preferable to further anti-oestrogen therapy. Therefore, detection of emergent *ESR1* mutations may be important in selecting the correct line of therapy for ER positive breast cancers, selecting non-oestrogen targeted therapies if *ESR1* mutations are present and alternative chemotherapeutic agents are still available.

In this study, all breast cancers patients were analysed in the same way and analysis did not take into account the molecular subtype. The subtypes of breast cancer most commonly used in the clinic are based on ER and HER2 status, splitting the cancers into four main groups to guide therapy (Cardoso et al., 2018). Genetic studies split breast cancer further into molecular subtypes luminal A and B, HER2, basal, claudin-low and a normal breast like group (Prat and Perou, 2011, Banerji et al., 2012) and more recently 10 genomic and transcriptionally defined subtypes (Russnes et al., 2017). The most important finding of these studies is that the mutational spectrum of these subtypes vary. For example, triple negative breast cancers (ER-/HER-) which includes basal breast cancers most frequently have *TP53* mutations but also have a range of other mutations in over 50 different genes (Shah et al., 2012), which makes selection of genes for targeted NGS analysis difficult. At

the time of diagnosis some triple negative breast cancers already display a wide variety of clonal evolution, such that tracking individual mutations in total cfDNA may be less informative. ER negative cancers are also more likely to have *TP53* mutations whilst ER positive patients have *PIK3CA* as the predominate mutation (Nik-Zainal et al., 2016). Patients with Germline *BRCA* gene mutations also have a different mutational signature (Nik-Zainal et al., 2016). Therefore, knowing the subtype of breast cancer may help inform which mutations to profile for ctDNA monitoring.

In total 5 different mutations in *TP53* and *PIK3CA* and 2 different *ESR1* mutations and were tracked in cfDNA. Although these would appear to be the most common mutations (Guttery et al., 2015, Shaw et al., 2017, O'Leary et al., 2018, Page et al., 2017, Fribbens et al., 2018) in ctDNA the small sample size suggests that only a proportion of the full range of ctDNA mutations are present in this cohort. Additional mutations in *TP53* and *PIK3CA* are well characterised as well as mutations in other genes including *AKT, GATA3, PTEN, KRAS, CDK2NA, CDH1 and CTNNB1* (Cohen et al., 2018, Fribbens et al., 2018, Garcia-Murillas et al., 2015). Additional *ESR1* mutations have also been detected in ctDNA including p.Y537N, p.Y537C, p.L536H and p.L536R (Fribbens et al., 2018, Schiavon et al., 2015). Together these suggest that the use of panels involving additional genes and gene regions may improve ctDNA detection.

3.4.5.2 Neo-adjuvant Cohort

Analysis of blood samples from 10 patients taken pre and post neoadjuvant chemotherapy showed 5 had detectable ctDNA and 2 of these patients relapsed, while none of 5 patients who were negative for ctDNA have relapsed at 1-3 years follow-up. Although a small cohort study, this firstly shows that the presence of ctDNA pre-surgery is a poor prognostic sign. Secondly, this shows that in patients receiving neoadjuvant chemotherapy this information could potentially be useful for selection of adjuvant therapy, by highlighting those patients at increased risk of relapse. The other patients with detectable plasma (CRB20, 50 and 104)

should be monitored closely after surgery as ctDNA data suggests that they may be at high risk of relapse.

If the relapse rate, and level of ctDNA positivity is as high as this exploratory cohort suggests, with 50% ctDNA positivity and a relapse rate of 40% in the ctDNA positive patients; a study of just 100 patients would yield 50 patients who were ctDNA positive and 20 relapses in the first 3 years of follow-up. In addition, patients would need to be followed post surgery with several blood samples, to find evidence for utility of ctDNA in detecting early relapse.

3.4.5.3 The merits of NGS profiling and further work

Interestingly, ctDNA variants were successfully detected in the neoadjuvant cohort using the Oncomine™ assay without the need for prior analysis of the tumour DNA, suggesting that primary tumour analysis is not always required. Results also show that the majority of patients have only 1 or 2 variants with this larger panel (100+ hotpsots) confirming the results in the palliative cohort and justifying the utility of a smaller more focussed NGS panel.

Increasing the number of genes covered by NGS sequencing either by WES or the use of panels such as the Oncomine™ breast panel, could improve detection of somatic tumour SNVs to above 90% of patients, dependent on tumour type (Liang et al., 2018, Nik-Zainal et al., 2016). Theoretically, this would increase the proportion of patients in whom ctDNA tracking would be possible, and is the reason why many larger panels are in development. This suggests that more patients had ctDNA than were shown in this study. This seems reasonable to assume, as within the palliative cohort a number of those patients without a detectable SNV in their tumour DNA went on to have *ESR1* mutations present in total cfDNA, indicating that ctDNA was present after all. Therefore, the use of larger sequencing panels or genome wide profiling, would appear to be more effective at ctDNA based patient monitoring and may provide a small improvement in ctDNA detection, if sensitivity of SNV detection in total cfDNA can be maintained. However, it remains unproven as to whether

the increase gene coverage of such panels, leads to a sufficient increase in the degree of ctDNA detection to justify their increased cost, compared to smaller focussed panels.

Finally, as part of further work, the analysis of all the blood and tumour samples collected from the 22 patients in the neoadjuvant cohort using the Oncomine™ panel would be of great interest, in particular profiling the post-operative bloods to look for evidence of minimal residual disease by molecular relapse.

4 Whole exome sequencing of serial plasma cfDNA captures tumour heterogeneity in a patient with metastatic breast cancer

4.1 Introduction

Although NGS profiling with targeted sequencing panels such as the AcquiRes panel and the Oncomine™ breast cfDNA panel can detect key driver mutations in tumour and plasma, these may only represent a subset of all mutations in the tumour. The use of these panels may therefore miss key driver events and potential therapeutic markers. In order to detect the full range of mutations WES and/or WGS is required. This generates a large amount sequencing data that requires a number of data processing steps or 'pipelines' to enable genome alignment and variant calling.

The recent development of total cfDNA WES potentially provides an alternative method for determining tumour genetics in plasma samples with sufficient levels of ctDNA, and may allow mutation tracking by following any important genetic event within the cancer (Toledo et al., 2018, Butler et al., 2015).

Butler et al. (2015) showed the clinical utility of using total cfDNA WES in two patients with metastatic disease; one with sarcoma and one with breast cancer. In both cases, a single blood sample was available with high levels of total cfDNA at 63 and 98 ng/ml respectively. Sequencing mean depth was 160-309. In the sarcoma patient, total cfDNA WES detected 47 of the 48 variants identified in the matched tumour DNA, including activating mutations in *KRAS* and *PIK3CA*. 15 additional mutations were identified in the total cfDNA, which had not initially been identified in the tumour DNA. 4 of these 15 mutations were subsequently demonstrated to be present in the primary tumour WES data but below the 10% VAF

threshold set for calling a tumour mutation. In the patient with breast cancer total cfDNA WES similarly identified 38 of 48 mutations present in tumour DNA. DNA from both the primary and a metastatic liver lesion were analysed. 17 additional mutations were identified in the total cfDNA, which had not initially been identified in the tumour DNA. 9 of these 17 mutations were subsequently demonstrated to be present in the primary tumour WES data. An activating *PIK3CA* p. H1047R mutation was detected in the primary lesion, but was absent from the metastatic lesion and matched cfDNA, where an *ESR1* p.D538G mutations was detected. Of note, the VAFs detected in total ctDNA correlated more closely with the VAFs in the metastatic liver lesion (sampled 2 weeks prior to the blood sample) than with the primary lesion from 50 weeks previously. This showed that total cfDNA WES was more correlated with the synchronous metastatic site than the original primary in the breast cancer patient, in keeping with the idea that total cfDNA provides a 'snapshot' of the tumour genetics. This study also showed that total cfDNA WES is able to identify clinically important mutations and at relatively low frequencies, as the mean VAF was 3.7% in the sarcoma case. They did however use 100ng of cfDNA for library preparation, from up to 25mls of plasma (50+mls of blood), which would be impractical for the majority of patient samples with lower levels of total cfDNA. In addition, patients may be unwilling to provide several large blood samples.

Toledo et al. (2018) have also recently demonstrated cfDNA WES in a patient with metastatic colorectal cancer. Their workflow used a lower amount of cfDNA (15ng total cfDNA), within the range of most metastatic patient samples. High concordance was also demonstrated between exome profiles from total cfDNA WES and tumour DNA from a liver metastasis. The cfDNA WES detected 68 of the 88 (77.3%) mutations detected in the liver metastasis, including driver mutations in *APC* and *TP53* and a putative driver mutation identified in *KDR* which encodes VEGFR2. This *KDR* variant was shown to be responsible for primary resistance to anti-VEGFR therapy, and was also confirmed to be present in subsequent blood samples. This study showed that total cfDNA WES libraries can be constructed from total cfDNA levels found in the majority of cancer patients, and the resulting libraries provide high concordance with synchronous tumour biopsies. This

suggests that total cfDNA WES could become a standard way to carry out global genomic profiling of advanced and progressing cancers where ctDNA levels are typically high, capturing both spatial and temporal tumoural heterogeneity through serial blood sampling. Ultimately this may allow doctors and scientists to track the evolution of the tumour genome without the need for serial tumour biopsies.

As a pilot study, this chapter follows one patient's journey with serial plasma exome analysis and relating results to their disease history.

4.2 Hypothesis, Aims and Objectives

4.2.1 Hypothesis

The hypothesis tested in this chapter is that total cfDNA WES of sequential blood samples will provide information on the evolution of a cancer over multiple courses of therapy from relapse until death. Such information may allow us to predict response and resistance to anticancer therapies and improve the outlook for such patients. This is the first time >2 follow-up cfDNA samples have been studied via WES.

4.2.2 Aim

The aim of this chapter was to compare the mutational profiles of the primary tumour, relapse biopsies and serial plasma samples, taken over a 2 year period in a patient with metastatic breast cancer, undergoing palliative therapy and relate findings to the clinical history of the patient. This is the first time multiple cfDNA samples have been analysed this way in a single patient.

4.2.3 Objectives

i. What is the concordance between tumour, relapse biopsy and serial samples?
ii. What proportion of detected variants are unique to tumour and cfDNA?
iii. Do changes in mutational profiles correlate with the patient's clinical course?

4.3 WES Sequencing

Patient CRB24 was selected as 10 blood samples were available along with primary and relapse tumour DNA samples. The patient was also known to have a *PIK3CA* p.E545K mutation present in ctDNA (see chapter 3). One buffy coat (lymphocyte DNA) sample, three tumour DNA samples, and total cfDNA from 8 of the 10 blood samples were selected for exome analysis (Table 17).

Samples were sent to a collaborator's laboratory (Dr Rodrigo Toledo, the University of Madrid) where libraries were prepared for WES using their cfDNA exome workflow (Toledo et al., 2018). An overview of the libraries can be seen in Figure 25. The size of the libraries generated from total cfDNA in lanes 5-12 appeared tighter in distribution than the tumour and lymphocyte samples in lanes 1-4. This may have been due to the highly fragmented nature of total cfDNA resulting in a smaller more consistent DNA fragment size, in comparison with DNA isolated form FFPE tissues. All libraries were considered suitable to proceed and sequencing was carried out using the Illumina HiSeq4000 platform. The resulting data (FASTQ files) were then shared with myself and Dr R Hastings (bioinformatics lead for the Shaw lab).

ID	Sample	Identifier	Sample Date	ng sent
1	T15/1	DNA from Curative Resection tumour sample	13/06/2005	80
2	T16/1	DNA from block 1 of Local breast relapse	13/01/2015	80
3	T16/2	DNA from block 2 of local breast relapse	13/01/2015	80
4	B14/15BC	Buffy coat / Lymphocyte DNA	10/03/2015	40
5	B14/15	Timepoint total cfDNA 1	10/03/2015	20
6	B31/15	Timepoint total cfDNA 2	21/04/2015	20
7	B90/15	Timepoint total cfDNA 4	15/09/2015	40
8	B14/16	Timepoint total cfDNA 5	26/01/2016	20
9	B44/16	Timepoint total cfDNA 6	19/04/2016	40
10	B91/16	Timepoint total cfDNA 7	19/07/2016	20
11	B143/16	Timepoint total cfDNA 8	09/11/2016	20
12	B20/17	Timepoint total cfDNA 10	16/03/2017	40

Table 17 - **Samples sent for whole exome analysis**.

a)

b)

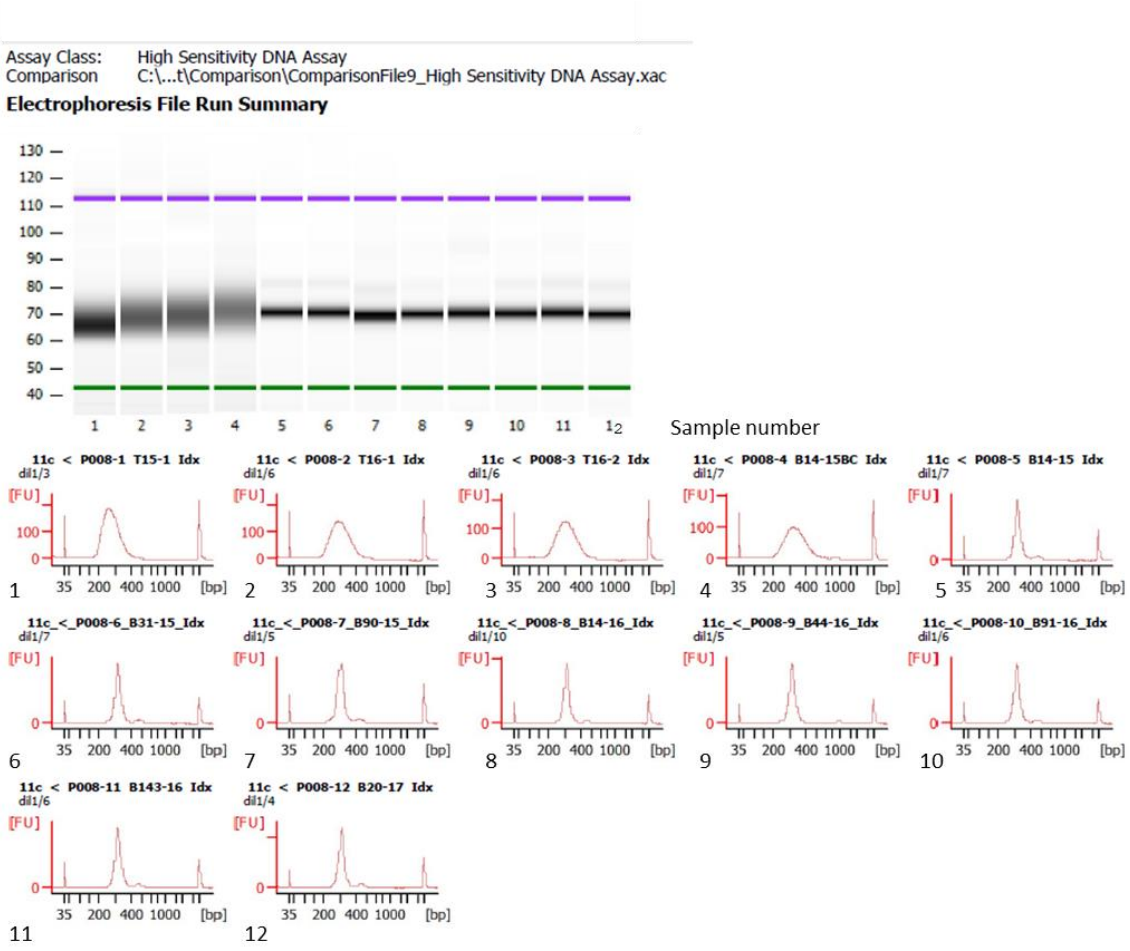


Figure 25 – Assessment of library quality by high sensitivity bioanalyser. a) electrophoretogram summary b) electrophoretogram 2-D. numbers 1 -12 relate to samples as per Table 17.

4.4 Mutation profiles of tissue and serial cfDNA samples

Dr Robert Hastings undertook the bioinformatics behind variant calling in this section. The FASTQ files were converted to aligned BAM files as described in chapter 2.16. Somatic mutations were called using Mutect2, version 3.6 (McKenna et al., 2010) with the buffy coat sample used as a reference to exclude germline mutations. 3097 somatic mutations called by Mutect2 were initially filtered using low stringency parameters (Table 18) approximating the parameters used by Butler et al. (2015).

Low Stringency Filter Parameters	
Single nucleotide variants	Deletions and insertions
• Depth >30	• Depth >50
• Mutant Read >5	• Mutant Read >10
• Frequency > 1%	• Frequency > 1%

Table 18 – **Low Stringency Filter Parameters**

812 mutations passed these low stringency filters (Figure 26). Tumour samples T15-1 and T16-2 had variants at a higher frequency than tumour sample T16-1, suggesting that these 2 samples had a higher tumour cell content than T16-1. Of the 8 serial cfDNA samples, B90-15 and B20-17 had mutations at a higher frequency than the other 6 samples.

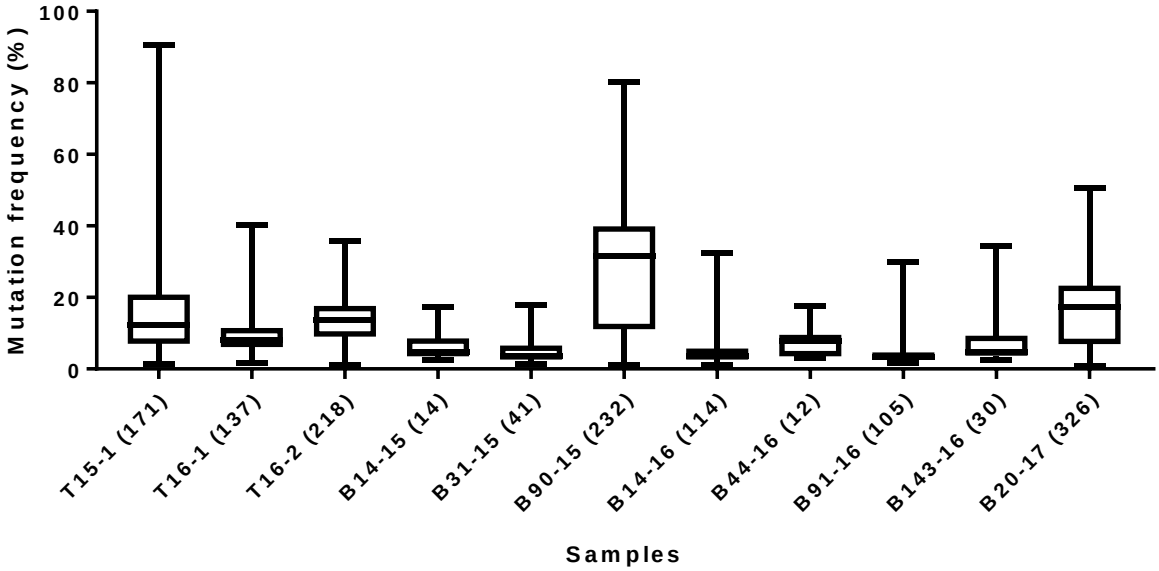


Figure 26 – Distribution of VAF for mutations in each sample that passed the low stringency filter, Number of mutations in each sample shown in brackets, bars show full range of mutation frequency, boxes show interquartile range.

Using the low stringency filter, a total of 382 mutations were present between tumour samples T15-1, T16-1 and T16-2. Of note only 27 of 382 mutations (7.1%) were shared between all three samples, with a further 90 (23.5%) shared between at least two samples (Figure 27). The majority of the variants detected (265 of 382, 69.3%) were private mutations not shared between any samples. Tumour DNA samples T16-1 and T16-2 had the most shared mutations (80) which is not surprising given that they originate from 2 separate regions biopsied from the same local relapse.

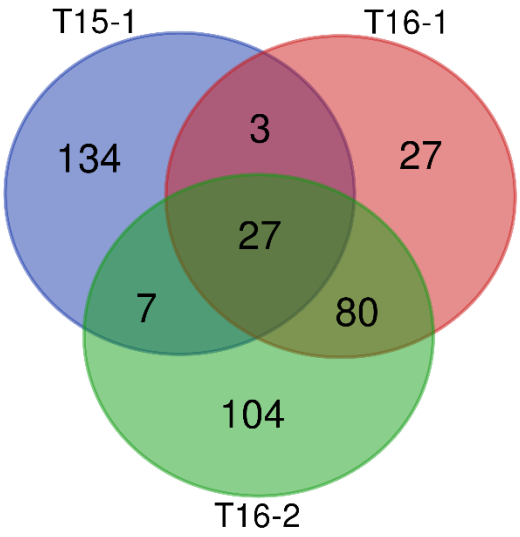


Figure 27 – **Overlap of mutations between the 3 tumour samples.** Using Low stringency filter.

Figure 28 shows the correlation of the VAF of shared (overlapping) mutations between samples using low stringency filtered data. T15-1 and T16-2 had the highest correlation of shared somatic mutations VAF between tumour samples. Correlation values between tumour and cfDNA samples were highest for B90-15 and B20-17, with lower levels of correlation for samples B14-16 and B91-16. The other four total cfDNA samples had less than 3 shared somatic mutations with the tumour DNA samples and are therefore not discussed further in this section.

When variants in the 3 tumour samples (T15-1, T16-1 and T16-2) and cfDNA samples B90-15 and B20-17 were compared, this showed that 23 of 639 mutations (3.6%) were shared between all samples (Figure 29). The 23 common mutations were reviewed for gene function using Entrez Gene (https://www.ncbi.nlm.nih.gov/), gene cards (https://www.genecards.org/) and the human protein atlas (https://www.proteinatlas.org). Three of the 23 genes (*, PP2R2D, PTDSS2, and MAP4K2*) had a putative role in cell signalling or cell proliferation, with one (*MAP4K2)* having a putative role in the MAP-kinase pathway, which is well characterised in cancer (Burotto et al., 2014). Overall sample T16-2 had more shared variants with the total cfDNA samples

than T16-1. Therefore, T16-2 was selected as the tumour relapse sample best suited for further analysis.

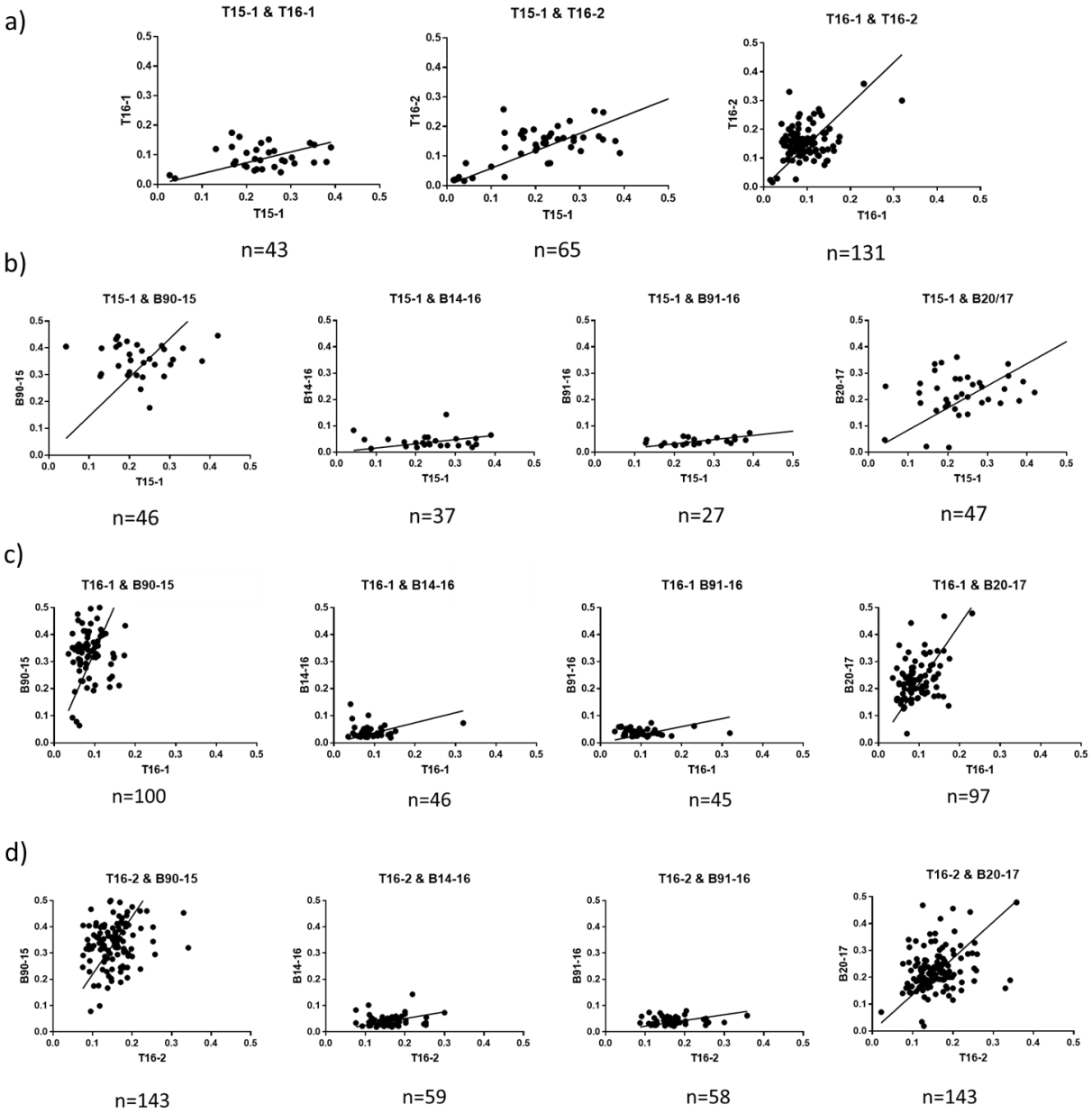
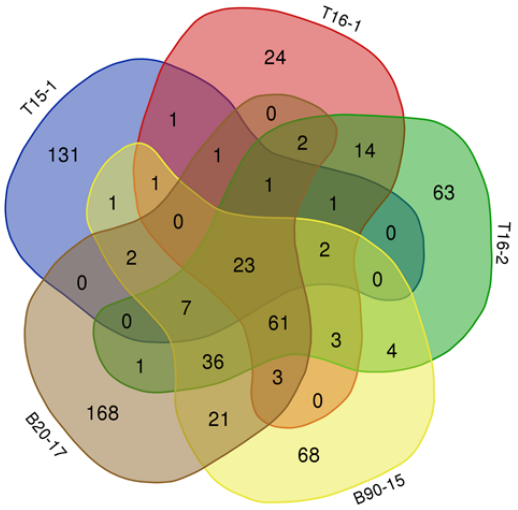


Figure 28 – **Scatterplots VAF results for shared mutations between tumour DNA samples and total cfDNA samples.** Variants filtered using low stringency parameters. Results shown for all samples that shared > 3 mutations. Number of shared mutations (n) shown for each pair of samples.

a)



b)

Gene	Location	Protein	Type mut.	T15-1	T16-1	T16-2 VAF (%)	B90-15	B20-17	Putative Gene Function
RPS6KA1	chr1:26881636:C/G	Intron	NA	39	12.5	11	52.9	26.8	Serine/Threonine kinase family protein Mediates phosphorylation of MAPK pathway mediates cell proliferation
SPATA6	chr1:48877241:C/T	T86T	Silent	16.7	12.7	10.8	40.4	33.5	Sperm Formation
KPRP	chr1:152733570:G/C	W502C	Missense	28	8.2	13	40.8	26.4	Keratinocyte protein
IGKV1D-8	chr2:90260082:C/T	P62L	Missense	26.3	5.9	16.1	33.8	25.7	Immunoglobulin light chain
IL36RN	chr2:113817033:G/T	A6A	Silent	20	6	13.9	31	18.2	Interleukin cytokine
XIRP2	chr2:168107351:A/T	Y3150F	Missense	20	10.7	11.9	37.6	18.6	cytoskeleton
MYLK	chr3:123457890:C/A	A148S	Missense	22.3	5.1	14.3	52.8	36.1	Cytoskeleton
KLHL6	chr3:183245702:C/T	T130T	Silent	35.2	7.4	15.6	57.6	33.5	B-lymphocyte antigen receptor signaling and germinal-center B-cell maturation
CCL28	chr5:43382070:T/C	K92K	Silent	25	15.2	20.2	62.1	28.5	antimicrobial gene chemotactic activity for resting CD4 or CD8 T cells
ADAMTS2	chr5:178585775:C/T	D361N	Missense	23.5	5.1	17.7	34.5	22.1	Metalloproteinase, cleaves type I +II collagen
SRSF3	chr6:36566709:G/C	G97A	Missense	28.6	7.8	16.3	29.4	24.8	mRNA splicing
LPA	chr6:161007601:T/C	M1337V	Missense	35.3	13.4	24.8	57.4	29	Lipoprotein A - Inhibits tissue-type plasminogen activator 1.
POU6F2	chr7:39379402:C/A	P225T	Missense	17.3	6.8	16.1	33.3	24.3	Probable Transcription factor likely to be involved in ganglion differentiation
CHAT	chr10:50835788:C/T	D238D	Silent	19.6	6.4	19	29.8	20	catalyzes the biosynthesis of the acetylcholine
PPP2R2D	chr10:133769286:C/T	P167L	Missense	17.1	6.9	18.6	44.3	15.8	B regulatory subunit of protein phosphatase 2A (PP2A) that plays a key role in cell cycle by controlling mitosis entry and exit oppose the action of kinases and phosphorylases and are involved in signal transduction.
PTDSS2	chr11:489618:C/G	L334V	Missense	21.9	8.6	16.2	41.2	27.9	The protein encoded by this gene catalyzes the conversion of phosphatidylethanolamine to phosphatidylserine, a structural membrane phospholipid that functions in cell signalling, and apoptosis
MAP4K2	chr11:64563782:T/A	I572F	Missense	13.1	12	12.9	39.9	18.7	Serine/threonine-protein kinase which acts as an essential component of the MAP kinase signal transduction pathway
ATG2A	chr11:64684560:G/A	V16V	Silent	30.2	9.1	11.7	33.8	20	Required for both autophagosome formation and regulation of lipid droplet morphology and dispersion
CDCA3	chr12:6959641:G/T	T80T	Silent	25	10.8	15.7	35.9	21	F-box-like protein which is required for entry into mitosis
RYR3	chr15:33603212:C/T	Not known		16.7	17.5	17.4	43.3	31.1	Calcium channel that mediates the release of Ca(2+) from the sarcoplasmic reticulum
C15orf54	chr15:39544403:A/T	No Protein		23.1	8.2	16.7	38.9	27.8	Putative RNA gene
DHX38	chr16:72142267:A/G	K1036E	Missense	34.3	14	16.7	59.7	24	Splicing
ELFN2	chr22:37770363:G/A	C404C	Silent	38	7.6	15.1	35.1	19.5	Expressed brain tissue

Figure 29 – **Shared mutations for the three tumour samples (T15-1, T16-1 and T16-2) and two cfDNA samples (B90-15 and B20-17)**. Using low stringency filter. a) Shared number of mutations b) The 23 mutations shared across all sample. Highlighted genes have a putative role in cell signalling or the cycle division.

In order to select only those mutations with the highest confidence, next a higher stringency filter was applied to the Mutect 2 output (Table 19), using the same number of mutant reads as described by Toledo et al. (2018).

High Stringency Filter Parameters	
Single nucleotide variants	Deletions and insertions
• Depth >30 • Mutant Read >20 • Frequency > 5%	• Depth >50 • Mutant Read >20 • Frequency > 5%

Table 19 – **High Stringency Filter Parameters**

As the most informative total cfDNA samples were B90-15 and B20-17 the mutations in these 2 samples were compared to T15-1 and T16-2.

Using the high stringency filter, 364 mutations were identified across all samples. A high proportion of the mutations present in the tumour DNA samples were also present in the total cfDNA samples (B90-15 and B20-17) which also had a high level of unique mutations (Figure 30).

195 mutations were identified in samples T15-1, T16-2 and B90-15. Of the 195 mutations, 65 (33.3%) were shared between the tumour and total cfDNA samples. Of the 102 mutations identified in the tumour DNA samples, 65 (63.7%) were also present in the cfDNA sample and only 1 of 10 mutations (10%) that was shared between the two tumour samples was not present in the total cfDNA (Figure 31a). 36 unique mutations were identified between the two tumour samples, and 93 unique mutations in the cfDNA sample.

234 mutations were identified in samples T15-1, T16-2 and B20-17. Of the 234 mutations, 57 (24.3%) were shared between the tumour DNA and total cfDNA samples. Of the 102 mutations identified in the tumour DNA samples, 57 (55.9%) were also present in the cfDNA and 2 of 10 mutations (20%) that were shared between the two tumour samples were not present in the total cfDNA sample (Figure 31b). 45 unique mutations were identified between the two tumour samples, and 133 unique mutations in the cfDNA sample.

222 mutations were identified in the two cfDNA samples B90-15 and B20-17. Of the 222 mutations 126 (56.8%) were shared between the two samples, 32 (14.4%) were unique to B90-15 and 64 (28.8%) were unique to B20-17 (Figure 33). 70 mutations were shared between the two total cfDNA samples that were not present in either of the two tumour samples (Figure 34a). The total cfDNA samples were more similar to each than with either tumour sample; B90-15 shared 65 of 158 mutations (41.1%) with the two tumour samples and B20-17 shared 57 of 190 mutations (30%) with the two tumour samples. This homology reduced further if the two tumour samples were treated separately; B90-15 shared 13 of 158 mutations (8.2%) with T15-1 and 61 of 158 mutations (38.6%) with T16-2, B20-17 shared 11 of 190 mutations (5.8%) with T15-1 and 54 of 190 (28.4%) with T16-2 (Figure 32, Figure 33, Figure 34a).

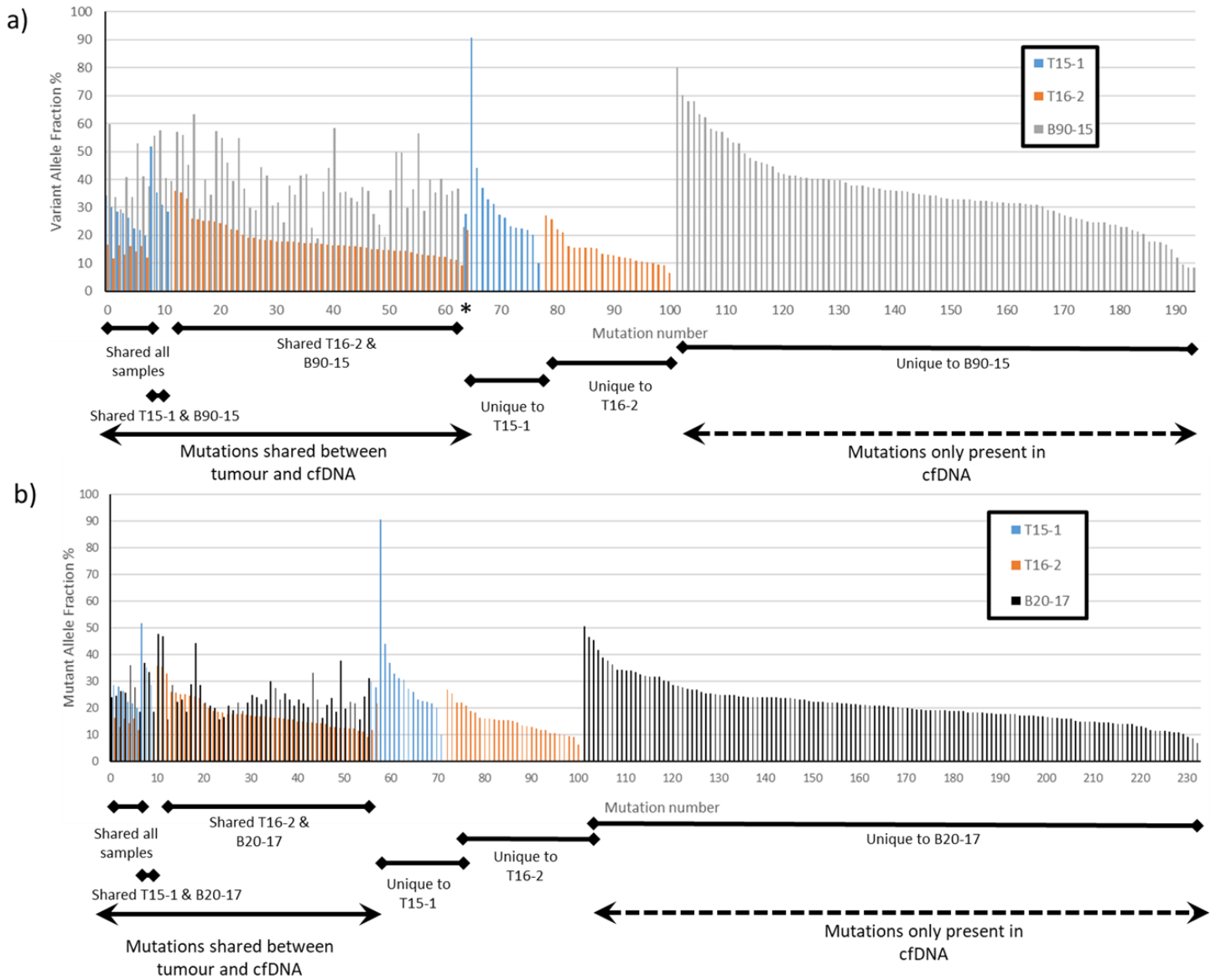


Figure 30 - **Overview of WES variant calls between tumour DNA samples T15-1 and T16-2 and total cfDNA samples.** a) B90-15 * shared T15-1 and T16-2 mutations b) B20-17. Variants filtered using high stringency parameters.

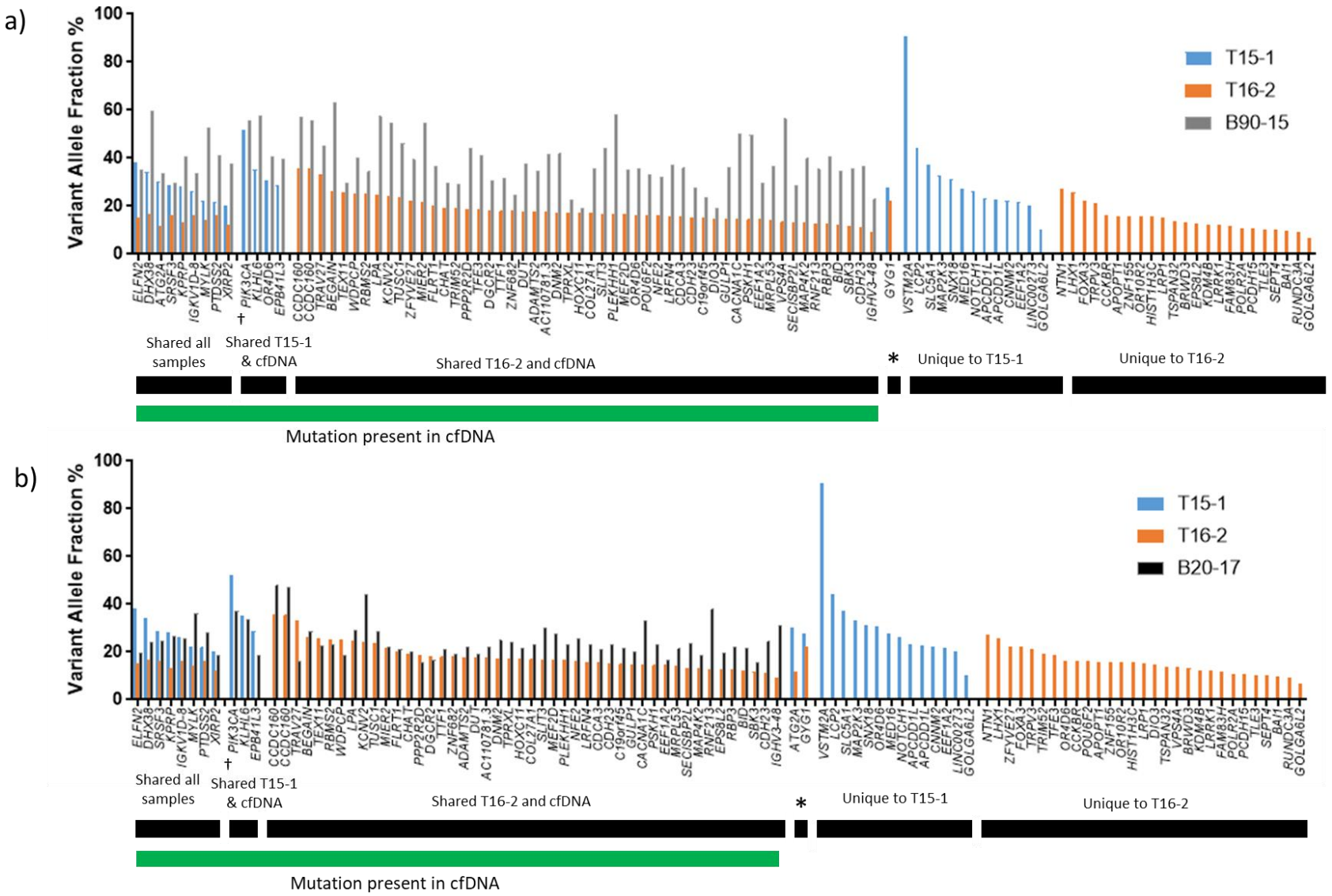


Figure 31 - **Comparison of WES variant calls between tumour DNA samples T15-1 and T16-2 and total cfDNA samples B90-15 and B20-17**. Variants filtered using high stringency parameters. a) Comparison with B90-15 b) Comparison with B20-17. *Mutations shared between tumour samples T15-1 and T16-2 only. †*PIK3CA* mutation previously identified from targeted NGS.

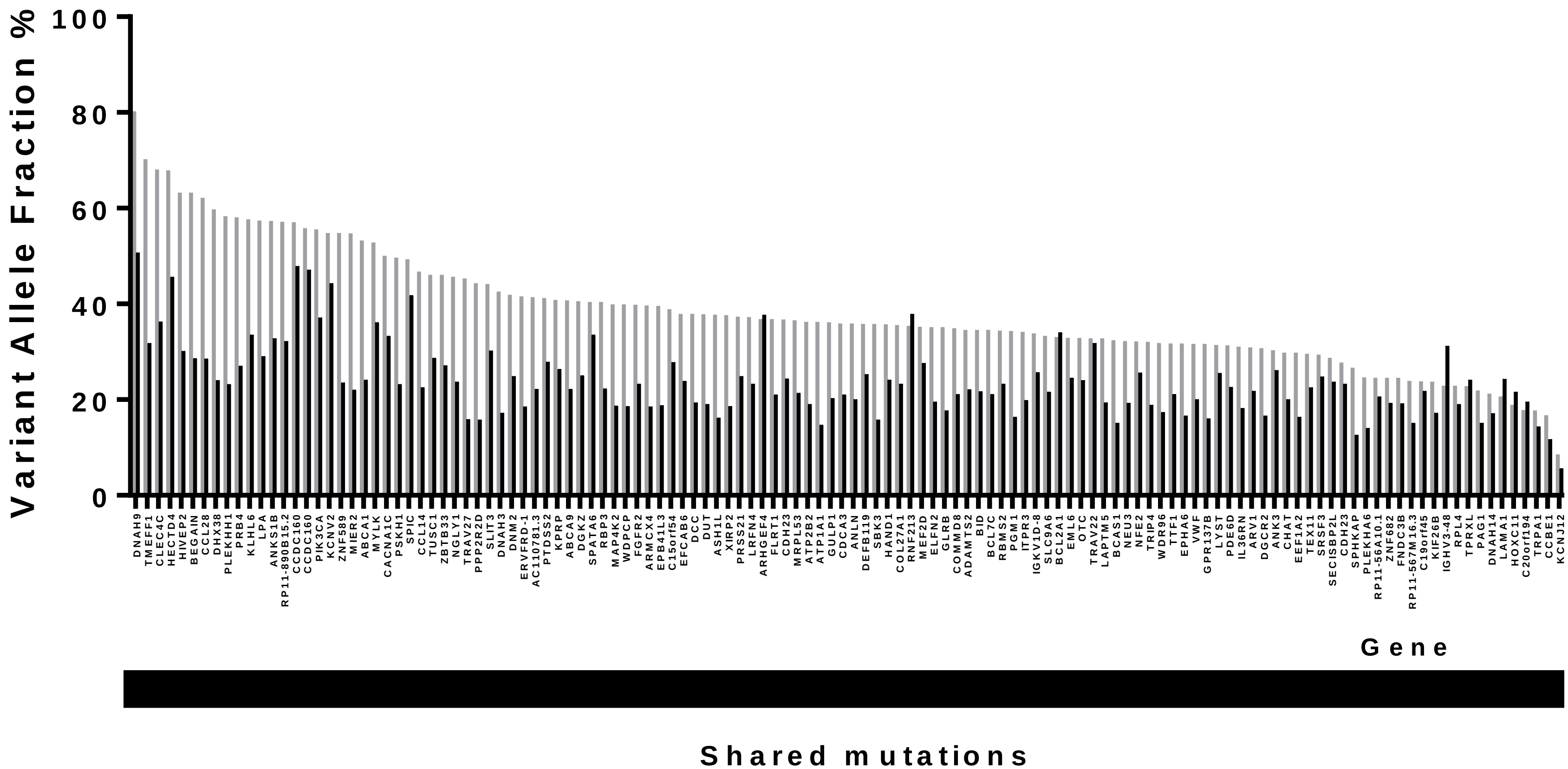


Figure 32 – **Shared WES mutations for B90-15 and B20-17.** Mutations filtered using the high stringency parameters

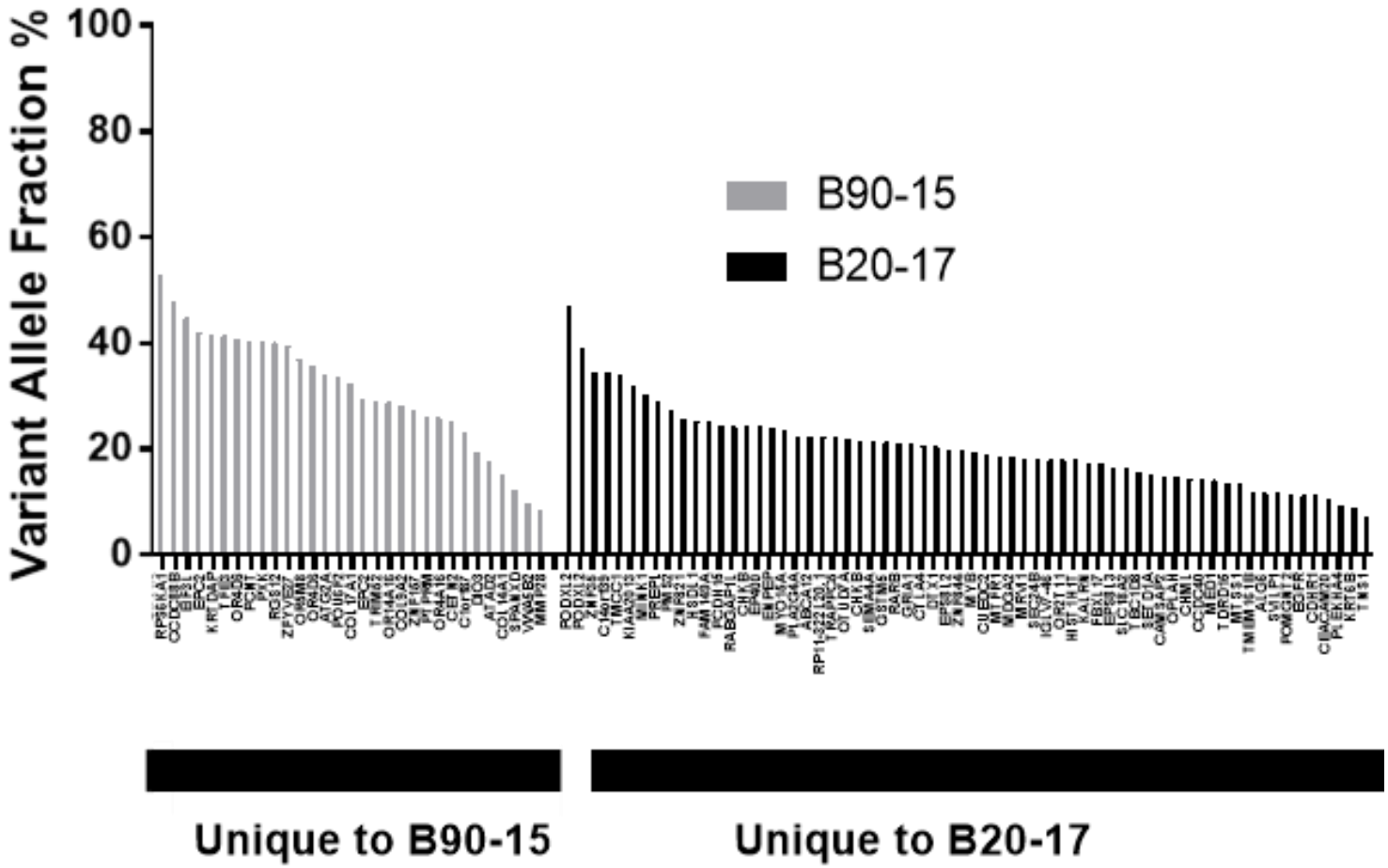
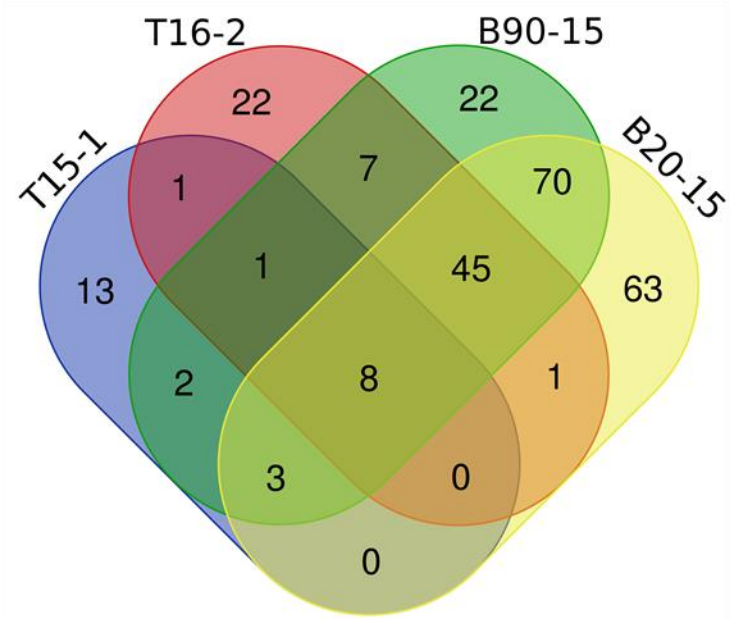


Figure 33 - **Unique mutations for B90-15 and B20-17.** Mutations filtered using the high stringency parameters.

Overall 8 mutations were present in all four samples (Figure 34b). This did not include the clonal *PIK3CA* p.E545K mutation which was absent from T16-2 by WES. However, this was detected previously by targeted NGS and ddPCR in the same sample. On further inspection the p.E545K mutation was detected in the exome data, with a mutant read depth of 19 (therefore excluded using the high stringency filter) and a VAF of 20.4% in the unfiltered data. One of the 5 previously identified cell division or cell signalling genes was still identified (*PTDSS2).* Of the other 7 genes, 2 (*XIRP2* and *MYLK)* had a role in cytoskeletal function, 2 (*SRSF3* and *DHX38*) had a role in spliceosome function, 1 (*IGKV1D-8*) had a role in immunoglobulin function, 1 *(KPRP*) in keratinocyte differentiation and 1 (*ELFN2*) was a silent mutation. Only a single gene (*XIRP2*) had a COSMIC ID mutation, identified in acute myeloid leukaemia. Why a mutation is *IGKV1D-8* should be present in the cfDNA but not in the paired buffy coat is not clear. It is a gene involved in the generation of the V region of the variable domain of the immunoglobulin kappa light chain (Entrez Gene https://www.ncbi.nlm.nih.gov/). It is possible that the mutation is the result of a clonal immunoglobulin rearrangement, potentially due to the previous chemotherapy but this has not been previously reported in cfDNA.

a)



b)

Gene	Protein	T15-1	T16-2	B90-15	B20-17	Type mutation	Protein	Function	Entrez Gene Gene Cards
		VAF							
KPRP	p.W502C	28.0	13.0	40.8	26.4	missense	p.W502C	Keratinocyte differentiation	Keratinocyte protein (skin)
IGKV1D-8	p.62L	26.3	16.1	33.8	25.7	missense	p.62L	Immunoglobulin function	Variable domain of immunoglobulin light chain
XIRP2	p.Y3150F	20.0	11.9	37.6	18.6	missense	p.Y3150F	Cytoskeleton	Expressed in muscle Interacts with F-actin COSM4001306
MYLK	p.A148S	22.3	14.3	52.8	36.1	missense	p.A148S	Cytoskeleton	Myosin Light chain kinase
SRSF3	p.G97A	28.6	16.3	29.4	24.8	missense	p.G97A	Spliceosome	mRNA splicing protein
PTDSS2	P.L334V	21.9	16.2	41.2	27.9	missense	P.L334V	Cell signalling and apoptosis	Catalyses the conversion of phosphatidylethanolamine to phosphatidylserine
DHX38	p.K1036E	34.3	16.7	59.7	24.0	missense	p.K1036E	Spliceosome	Alteration of RNA secondary structure
ELFN2	p.C404C	38	15.1	35.1	19.5	silent	p.C404C	No Data	Expressed in brain tissue

Figure 34 - **Overview of mutations present in tumour (T15-1 & T16-2) and total cfDNA (B90-15 & B20-17) samples**. Using the high stringency filter. a) Summary of shared mutations b) Putative function of the 8 variants common to all samples.

To investigate the degree of tumour evolution the unfiltered mutations were analysed using PHYLIP (http://evolution.genetics.washington.edu/phylip.html) by Dr Hastings. This programme generated a phylogenetic tree using the discrete character parsimony method. Samples B90-15 and B20-17 were more closely related to each other than to any other samples (Figure 35). They formed a clade together with the tumour DNA samples, with T16-2 being the tumour sample that B90-15 and B20-17 are most closely related to. B90-15 and

B20-17 are therefore most likely to be clonally derived from sample T16-2, and by logically extension the active metastatic disease in this patient is also derived from T16-2. The other total cfDNA samples were more distantly placed on the phylogenetic tree (not shown). These samples contained relatively few mutations overall (Figure 26).

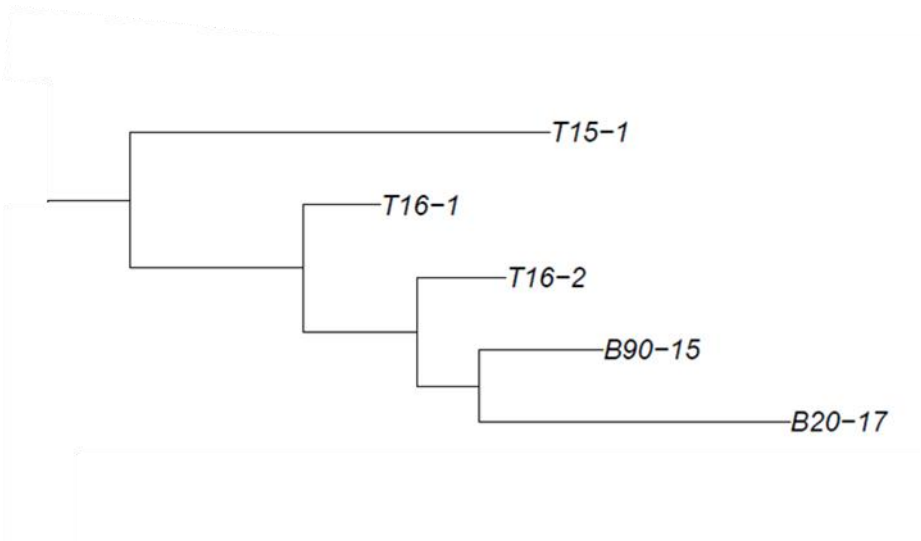


Figure 35 – **Phylogenetic tree of the WES samples**. Only SNV variants were used to generate the phylogenetic tree.

4.5 Mutational signatures

For each WES sample the mutational signature (MS) was generated from the unfiltered mutations using the deconstructSigs programme by Dr Hastings (Figure 36a&b). MS-3, the signature linked with deficient homologous recombination repair in *BRCA* mutation carriers, (Alexandrov et al., 2013) was the predominant signature in the primary tumour (T15-1), whilst MS-3 and MS-4 (linked with tobacco chewing) were the predominant profiles in the relapse biopsies (T16-1 and T16-2). The initial blood sample (B14-15) had a mixed profile and overall few detected mutations. MS-3 was detected in sample B31-15, which increased in frequency in samples B90-15 and B14-16. It then disappeared at time point B44/16, reappeared at B91-16 before then disappearing and reappearing at time points B143/16 and B20-17 respectively.

The blood samples taken at progression (B90-15, B14-16, B91-16 and B20-17) all had MS-3 as the predominant signature (Figure 36a-c) and in 2 of the 3 time points where a clinical response was seen (B44-16 and B143-16) the MS-3 signature disappeared. The predominance of MS-3 appears to track disease progression and provides additional information compared to monitoring *PIK3CA* p.E545K or other mutations alone; between samples B90-15 and B14-16, when clinical progression occurs, MS-3 increased in frequency but the *PIK3CA* mutation frequency decreased.

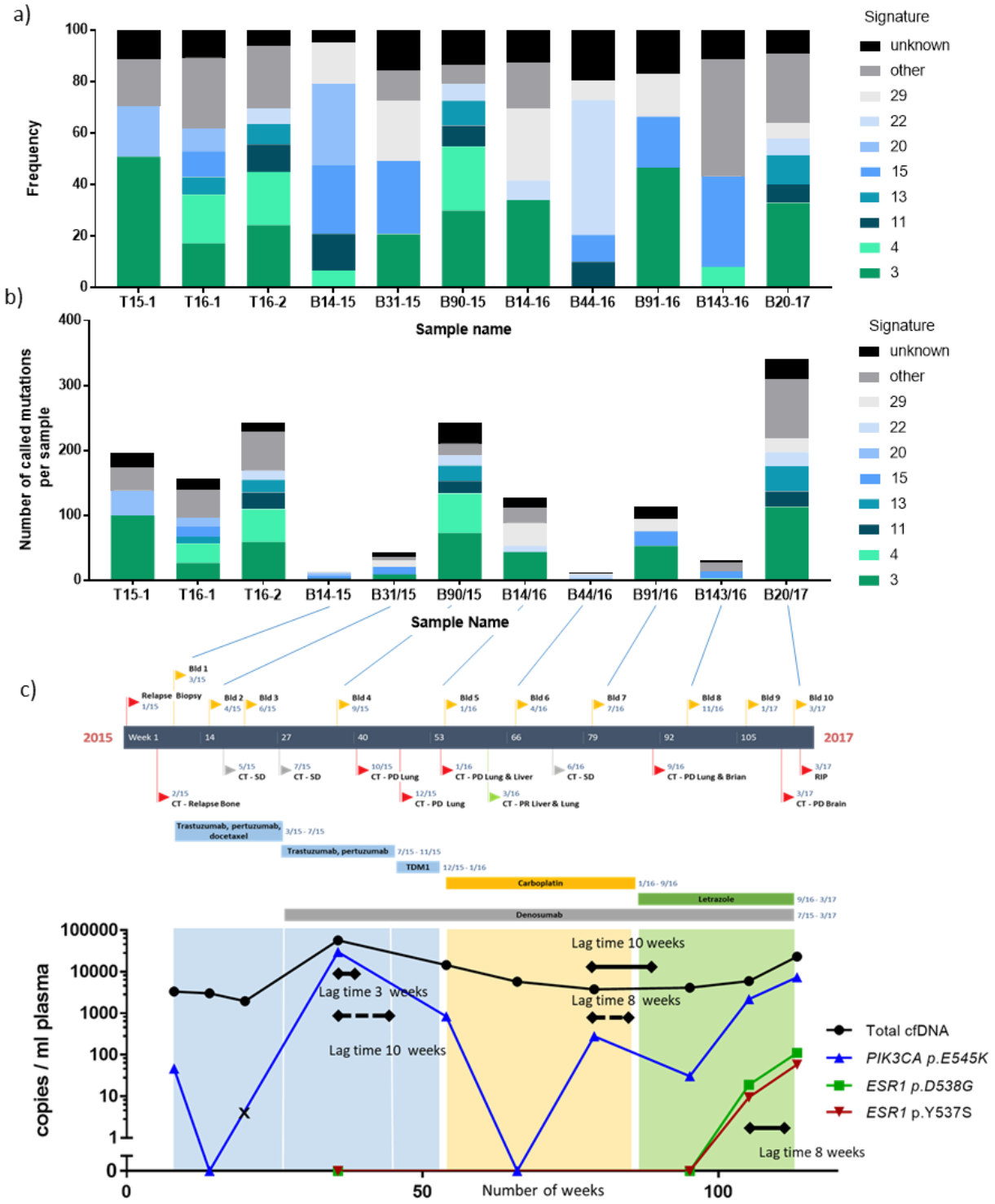


Figure 36 – **Mutation signature profile of WES samples.** a) Mutational signature profile of each sample. b) Mutational signatures profile as a proportion of the number of mutations (low stringency filter) detected in each sample. c) Patient Timeline from chapter 3.

4.6 Driver mutations

The unfiltered mutations were analysed using the Cancer Genome Interpreter (CGI) programme, available at https://www.cancergenomeinterpreter.org/home. This programme detects validated oncogenic alterations and predicts cancer driver mutations amongst mutations of unknown significance. Flagged genes are denoted as tier 1 or tier 2. Tier 1 mutations have documented evidence of activity that may drive cancer, as well as evidence that changes to protein activity promotes oncogenic transformation. Tier 2 mutations have strong evidence for playing a role in cancer but this is less comprehensive than for tier 1.

Four genes were flagged as containing potential driver mutations (Table 20). The tier 1 mutation, *PIK3CA* p.E545K, had previously been identified using targeted sequencing in chapter 3 and was present in the primary sample (T15-1), the relapse biopsies (T16-1 and T16-2) and 4 of the 8 total cfDNA samples. Three additional tier 2 driver mutations were identified; *DNM2* p.G146D*, DCC* p.E495K and *ATP1A1* p.W105X. All 3 tier 2 mutations were present in the relapse biopsies (T16-1 and T16-2) and in 2-5 of the total cfDNA samples but not in the primary tumour sample (T15-1).

Gene	Tier	T15-1	T16-1	T16-2	B14-15	B31-15	B90-15	B14-16	B44-16	B91-16	B143-16	B20-17
PIK3CA p.E545K	1	51.9	3.1	20.4	0.7	0	55.5	0	0	8	0	37.1
DNM2 p.G146D	2	0	11.5	17.3	0	0	41.9	0	0	2.2	0.4	24.9
DCC p.E495K	2	0	8	15.7	0	0	37.9	0	0	0	0	19.4
ATP1A1 p.W105X	2	0	8.1	10.9	0.5	0	36.2	0	0	2.3	0.5	14.7

Table 20 – **Flagged driver mutations from unfiltered WES data using CGI.** Frequency of mutation in each sample shown as VAF (%). Grey boxes indicate values not called by the programme and were recovered manually.

Of note none of these mutations were amongst the 23/8 mutations that were shared between T15-1,T16-2,B90-15 and B20-17 following the low or high stringency filters respectively (Figure 29 page 128 & Figure 34, page 136).

4.7 SCNA profiling

In addition to profiling variants in the exome sequence, SCNAs were identified using the allelic copy number caller Sequenza (Favero et al., 2015), following the workflow for BAM files in the R Vignette (https://cran.r-project.org/web/packages/sequenza/index.html) by Dr Hastings.

Analysis showed that SCNAs could be identified in those samples previously shown to have high mutational frequency (Figure 26) i.e. T15-1, T16-1 and T16-2, B90-15, B20-17 with amplification detected in chromosomes 1, 3 and 8 (Figure 37). SCNAs could not be detected in the other samples with lower mutational burden (See appendix 7.4) and associated lower cellularity (Appendix 7.5, Table 48) and ctDNA levels.

There was uncertainty concerning the HER2 status of the patient. IHC of the primary tumour (T15-1) was reported as HER2 3+, whereas the relapse biopsy, from which tumour samples T16-1 and T16-2 were derived, was reported as HER2 2+ and FISH negative. Review of the WES showed no amplification in the region encoding HER2 (17q12, chr17:37856254 – 37884915) in any sample including T15-1, suggesting a HER2 negative status.

In order to confirm this finding the DNA samples were also analysed for *ERBB2* gene amplification by ddPCR using *RPPH1* as the reference gene assay. The tumour biopsies had a copy number (CN) of <3, and 5 of the 8 total cfDNA samples had a CN >3, with B90-15 having a CN of 9.7 (Table 21), suggesting gene amplification. On further analysis of the exome data there was a large q-arm deletion of chromosome 14 in sample B90-15 (14q22.3-q32.31, chr14:57511686-102904545) and it was therefore hypothesised that the elevated copy number of *ERBB2* found via ddPCR in the total cfDNA samples could be an artefact of a *RPPH1* gene deletion (Figure 37). To test this theory sample B90-15 was reanalysed by ddPCR using two further reference genes; *GAPDH* and *CNTNAP1*, showing a CN of 3.8 and 1.6 respectively (Table 21). This shows that the CN obtained from *ERBB2:RPPH1* analysis was likely a result of reference gene deletion due to chromosomal instability within the

tumour, and the patient therefore is not *ERBB2* amplified. Clinically, a non-*ERBB2* amplified and therefore HER2 negative status for patient CRB24 is in keeping with the lack of clinical response seen to trastuzumab/pertuzumab and TDM-1 (Chapter 3, Figure 16, page 99).

Reference Gene	T15-1	T16-1	T16-2	B14-15	B31-15	B90-15	B14-16	B44-16	B91-16	B143-16	B20-17
ERBB2:RPPH1	2.35	1.80	1.92	2.70	-	9.7	3.67	-	3.51	3.17	3.03
ERBB2:GAPDH	-	-	-	-	-	3.8	-	-	-	-	-
ERBB2:CNTNAP1	-	-	-	-	-	1.6	-	-	-	-	-

Table 21 – **CN calculation by ddPCR for ERBB2 against 3 reference genes for DNA samples from patient CRB24**

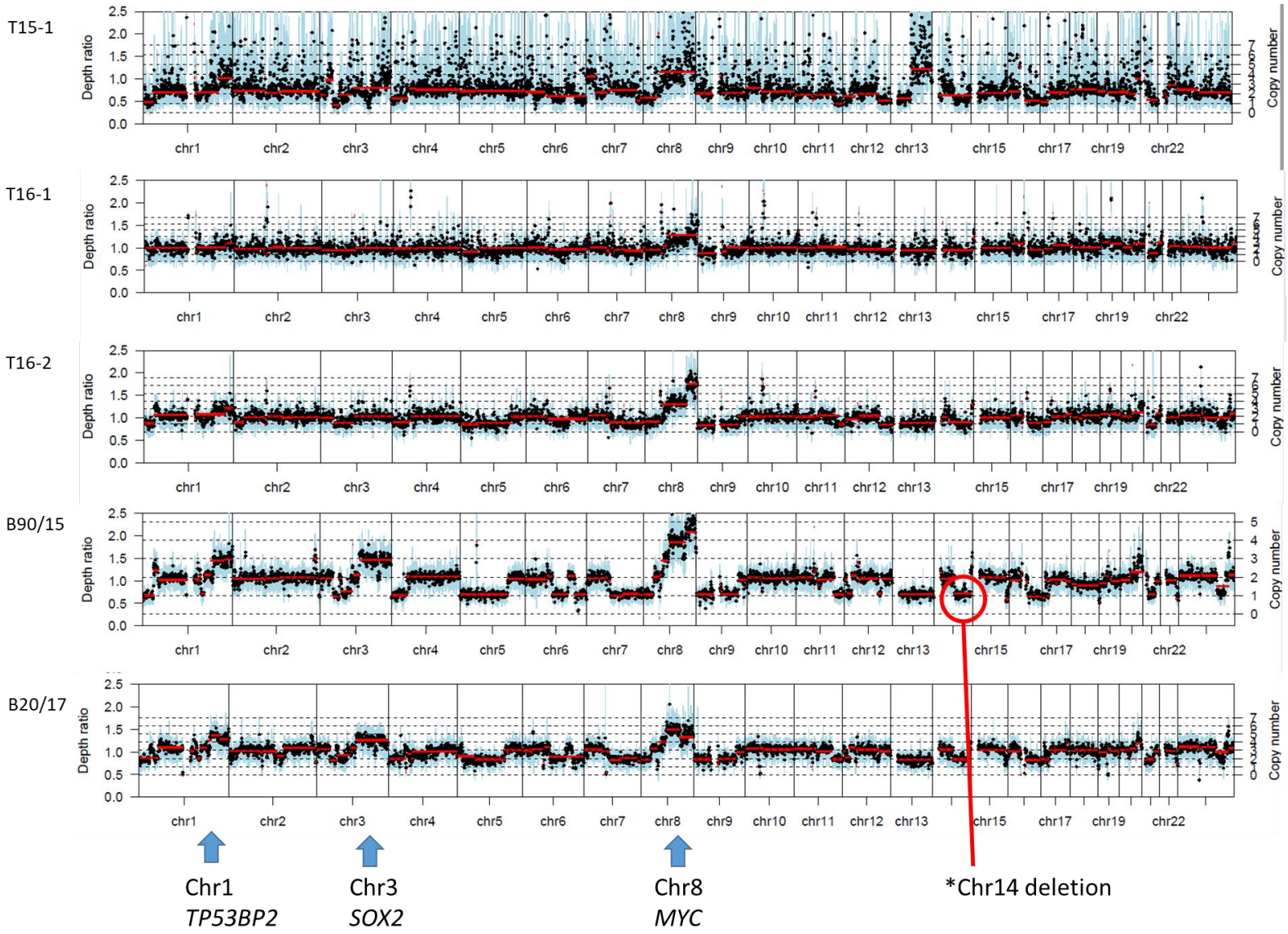


Figure 37 - **Whole exome analysis of SCNAs for samples with high mutational burden.** Arrows donate three points of amplification (CN ≥ 3) in chromosome 1, 3 and 8, and the putative cancer associated gene in the region. *Area of deletion of chromosome 14 in sample B90-15.

The highest level of amplification was in the q arm of chromosome 8, detected in all 3 tumour samples and both total cfDNA samples (Figure 37). The highest region of amplification with a CN of 4-16 in the five samples and spanning an approximate region in band 8q24.3 (chr8:145665359-145993019) contained 13 genes (Table 22a). None of these genes were known cancer related genes (analysed through CGI), but *GPT* was identified and has a defined role in the metabolism of glucose and amino-acids, and metabolism dysregulation is a hall mark of cancer (Table 22c).

a)

Sample	Chr	Region start	Region end	Copy number
T15-1	8	46868544	146171437	4
T16-1	8	146003327	146171437	4
T16-2	8	145665359	145993019	16
B90-15	8	145665359	145993019	5
B20-17	8	145665359	146000095	11

b)

Sample	Chr	Region start	Region end	Copy number
T15-1	8	118914377	145662113	6
T16-1	8	59324162	145662113	4
T16-2	8	118914377	145662113	6
B90-15	8	118185041	145662113	4
B20-17	8	103841709	145662113	5

c)

Gene	Name	Function	Information
TONSL	tonsoku like, DNA repair protein	Transcription regulation	Negative regulator NF-kappa B mediated transcription
KIFC2	kinesin family member C2		ATPase activity and microtubule motor activity
FOXH1	forkhead box H		Binds SMAD2 and activates activing response
GPT	glutamic--pyruvic transaminase	Metabolism	This enzyme catalyses the reversible transamination between alanine and 2-oxoglutarate to generate pyruvate and glutamate plays a key role in the intermediary metabolism of glucose and amino acids.
RECQL4	RecQ like helicase 4	DNA helicase	none
LRRC14	leucine rich repeat containing 14		none
CYHR1	cysteine and histidine rich 1		none
PPP1R16A	protein phosphatase 1 regulatory subunit 16A	cytoskeleton	Controls the phosphorylation states of regulatory myosin light chains
MFSD3	major facilitator superfamily domain containing 3		none
ZNF251	zinc finger protein 251		Unknown – zinc finger protein
ARHGAP39	Rho GTPase activating protein 39		Rho GTPase activating protein – intracellular protein expressed in all tissue
LRRC24	leucine rich repeat containing 24		Increased expression in thyroid and prostate cancer
C8orf82	chromosome 8 open reading frame 82		none

Table 22 – **Regions of amplification and gene function of chromosome 8.** a) Most amplified region b) amplified region containing *MYC* c) Putative gene function of genes in region a)

A second region of amplification in close approximation to the area of highest amplification was large (8q24.11-q24.3, chr8:118974377-145662133) and had a CN of 4-6 in the samples (Table 22b). The area of amplification in chromosome 8 contained 227 genes including the proto-oncogene *MYC*. The genes were run through CGI and only *MYC* was identified as a driver, the others being predicted passenger SCNAs. To confirm the amplification detected in WES analysis, ddPCR for *MYC* was completed using the reference gene *CNTNAP1*, as this

reference gene gave the closest result to diploid during *ERBB2* analysis. The results, which included two additional areas from the primary resection (T15-2 and a lymph node T15-3LN) showed detection of *MYC* amplification (defined as CN >3) in the tumour samples and B90-15 (Table 23). B14-16 had lack of amplification on ddPCR and WES. B20-17 was not amplified on ddPCR analysis but was amplified for the region containing *MYC* according to WES. Additional analysis of WES showed the ploidy level for B20-17 to be approximately 3, and 2 for all other samples (Appendix 7.5). This suggested that whole genome amplification had occurred in the tumour(s) producing the ctDNA in sample B20-17. Such an increase in ploidy level, would increase the amount of DNA from the reference gene in the subsequent ddPCR reaction and therefore would be predicted to reduce the calculated CN for *MYC.* Given that the *MYC* amplified samples all had a low calculated CN of <4, the increase in ploidy levels may be responsible for the *MYC* amplification no longer being detectable by ddPCR in sample B20-17 whilst still being detected by WES.

Sample	Type	CN	Amplified on ddPCR	Amplified on WES
T15-1	Initial surgery area 1	3.0	YES	YES
T15-2	Initial surgery area 2	3.3	YES	Not done
T15-3LN	Initial surgery lymph node	3.5	YES	Not done
T16-1	Relapse Biopsy 1	2.7	YES	YES
T16-2	Relapse Biopsy 2	3.6	YES	YES
B90-15	cfDNA 4	3.0	YES	YES
B14-16	cfDNA 5	1.9	NO	NO
B20/17	cfDNA 10	2.3	NO	YES

Table 23 - ***MYC* amplification via WES and ddPCR analysis**

Chromosome 3 had a large amplified region identified in the primary tumour sample (T15-1) and total cfDNA samples B90-15 and B20-17 and encompassed the region 3q13.13-q29 (Chr3:109052732-1097847031). The region had a CN of 3-4 and contained 728 genes. *SOX2* was identified as a predicted important amplification in this region following analysis with CGI, with the other genes predicted to be passenger amplifications.

Chromosome 1 had variable regions of amplification, present in all samples. The region 1q41-q42.11 (chr1:229366646-224202533) was amplified in samples T16-2, B90-15 and

B20-17 with a CN of 7-8. Three genes were present in this region, two pseudogenes and one protein encoding gene; TP53 binding protein *TP53BP2*, which is known to be amplified in some breast cancers (Lips et al., 2015). The region 1q21.3, was also amplified in sample B20-17 with a CN of 3 (chr1:142713152-161376557), and has been reported to be frequently present in metastatic breast cancer (Goh et al., 2017)

4.8 Discussion : WES of total cfDNA

WES detected a large number of mutations in primary and relapse tumour DNA samples and in total cfDNA samples. 3097 unique mutations were detected by the Mutect2 variant caller with an average read depth of 133. Of these 3097 mutations, 813 (26.2%) passed the low stringency filter and 268 (8.6%) passed the high stringency one.

Using the low stringency filter, significant numbers of shared mutations were detected between the primary samples T15-1 and the relapse samples T16-1 and T16-2 (Figure 27, page 126). As such there was no doubt that this was a case of relapsed cancer and not a new primary breast cancer. It was also possible to determine that of the samples from the relapse tumour, T16-2 shared more mutations with the cfDNA samples than T16-1 (Figure 29, page 128). As such, sample T16-2 was predicted to be the tumour region from which metastatic disease was clonally derived, and this was also supported by phylogenetic analysis. The TRACERx study on lung cancer (Abbosh et al., 2017) has also shown the above pattern, with some total cfDNA samples reflecting only a proportion of the mutations present in the primary samples.

Samples B90-15 and B20-17 had a higher mutational burden than the other total cfDNA samples (Figure 26) and also shared more variants with the tumour samples and were therefore the most useful samples to analyse with a high stringency filter (Figure 29, page 128).

Using the high stringency filter, the mutational profiles of the total cfDNA samples B90-15 and B20-17 were more similar to each other than to the tumour samples T15-1 and T16-2 (Figure 30-Figure 33, pages 131-134). These similarities between the two cfDNA samples confirms the reliability of the whole exome sequencing process and of variant calling and filtering. Many of the mutations detected were of higher frequency in the total cfDNA samples than in the tumour DNA samples, showing total cfDNA WES can be used to interrogate tumour genetics.

The total cfDNA samples (B90-15 and B20-17) were more similar to the relapse biopsy T16-2 than the primary biopsy T15-1 in terms of mutational profiling (Figure 34, page 136a). This is consistent with a process of tumour evolution between the time of the primary tumour and the subsequent relapse, which was a period of 10 years. Sample B20-17 was taken 78 weeks (548 days,) after sample B90-15 and it is apparent that the similarities in mutation profile between the tumours and the total cfDNA samples reduced over time and that B20-17 had more unique mutations. This suggests that the process of tumour evolution was ongoing and was responsible for the divergent mutational profiles (Figure 33, page 134)

Phylogenetic analysis of the mutations supported the evolutionary relationship between the tumour samples and B90-15 and B20-17 (Figure 35, page 137). The other total cfDNA samples clustered separately on the phylogenetic tree to these samples, which was likely an artefact of their low number of detected mutations.

The patient studied in this chapter had a germline *BRCA2* exon 14-16 mutation. Mutational signature (MS) analysis revealed that MS-3, the signature linked with *BRCA* mutation and associated with failure of DNA double-strand break-repair by homologous recombination (Alexandrov et al., 2013), could be found in all three tumour biopsy samples and 5 of the 8 total cfDNA samples (Figure 36, page 139). The frequency of MS-3 varied with response to treatment and was undetectable when the disease was well controlled. MS-3 also became more frequent in blood B14-16 at a time of disease progression when the main clonal mutation *PIK3CA* p.E545K became less frequent, suggesting it was disease with the MS-3

signature that was resulting in disease progression. There was also evidence in the WES MS profile that carboplatin, a chemotherapy specifically shown to be active in patients with MS-3 BRCA-related disease (Yamamoto et al., 2014), was the correct chemotherapy to give following blood B14-16, as shown by the loss of the MS-3 profile in the subsequent sample B44-16. Overall, no treatment permanently eradicated the MS-3 signature which was the predominant MS just before death, and therefore is likely the MS of the lethal clone.

The 8 mutations shared across all samples with the high stringency filter (Figure 34, page 136b) picked up a number of structural genes, which were unlikely to be driving mutations, and additionally identified *PTDSS2* which has a putative role in cell signalling and apoptosis but no identified role in tumorigenesis. Analysis with CGI picked up the tier 1 *PIK3CA* p.E545K mutation that had previously been identified with targeted NGS (Table 20, page 140). Due to the lower depth of reads compared to targeted NGS, WES / CGI detected the *PIK3CA* mutation in 5 of the 11 samples (45%) analysed in this chapter, although on review it was present in a further 2 samples at a frequency of <3%. In contrast by targeted NGS the mutation was detected in 9 of the 11 samples (81%).

Three further tier 2 mutations were specific to relapse tumour biopsies and total cfDNA samples, indicative of cancer evolution (Table 20, page 140). *DNM2* is a gene linked with cell migration, metastasis, and receptor endocytosis (Joshi et al., 2014, Xu et al., 2014). It may also have a role in regulation of hypoxia inducible factors (Joshi et al., 2014) and has been shown to have an increased expression in prostate cancer cells (Xu et al., 2014). *DCC* is frequently deleted or its expression reduced in breast cancer, glioblastomas, colorectal, gastro-intestinal, pancreatic and prostatic tumours (Koren et al., 2003) (Hsu and Shaw, 2000) Kataoka et al., 1995). *ATP1A1* shows decreased expression in human renal cell carcinomas than in the adjacent non-tumor tissues and it was therefore proposed that *ATP1A1* is a potential novel suppressor protein for renal cancer (Zhang et al., 2017). Therefore, all three tier 2 genes and *PTDSS2* may be evaluated for therapeutic intervention in the future.

The two emergent *ESR1* mutations were not detected by the CGI, and were not present on manual review of the unfiltered called variants. This is because both had a frequency of <1% which cannot be detected with an average depth of 133, as both would be predicted to have one or fewer reads at such depth.

A similar issue is apparent in some of the samples with Tier 1 and 2 genes (Table 20, page 140). Those samples that contain overall lower amounts of ctDNA, as defined by fewer called mutations, have driver mutation reads that do not get called by the CGI. Manual review of the unfiltered variants for these mutations revealed them to be present at low levels (2-8%). Automatic variant detection relies on a minimum number of reads of the mutation that can only be achieved in these samples with higher depth. This is difficult to achieve with WES. Samples which have a high fraction of ctDNA (such as B20-17 and B90-15) have more mutation reads at a set read depth, and therefore it is more easy to detect mutations in these type of samples.

Copy number analysis revealed that this patient was non-*ERBB2* amplified, but had amplification in the proto-oncogene *MYC* on chromosome 8 *(*section 4.7 page 141*).* Many potential drugs are under development for the treatment of patients with *MYC* amplification but none have yet reached widespread clinical use (Chen et al., 2018). *SOX2* was also identified as gene of interest, located in an amplified region on chromosome 3. *SOX2* has critical roles during mammalian embryogenesis has been implicated in growth, tumorigenicity, drug resistance, and metastasis in at least 25 different cancers, including breast cancer (Chen et al., 2008, Wuebben and Rizzino, 2017). Analysis of regions of amplification also identified amplified genes that may play a role in tumour growth and survival. Firstly, *GPT* which has a role in metabolism regulation, and therefore may play a role in the 'deregulation of cellular energetics' an emerging hallmark of cancer (Hanahan and Weinberg, 2011) and secondly *TP53BP2*, which binds the well characterised tumour suppressor *TP53* and therefore is likely to play a role in the 'sustaining proliferative signalling' hallmark of cancer*,* and is known to be amplified in some breast cancers (Lips et al., 2015). Finally, the region 1q21.3 was identified as amplified in sample B20-17, this has

previously been detected in total cfDNA in breast cancer patients and results in a S100A7/8/9–IRAK1 feedback loop which may be targeted by the small-molecule kinase inhibitor pacritinib (Goh et al., 2017).

4.8.1 Summary

In summary WES of total cfDNA is possible and provides useful information regarding the patient's tumour. Two total cfDNA (B90-15 and B20-17) had high mutational burden indicating high levels of ctDNA. These samples shared a significant number of mutations with the relapse biopsy T16-2 and less so with the primary biopsy T15-1 indicating that significant tumour evolution had taken place since the primary diagnosis. The results show total cfDNA WES can successfully probe the tumour genetics, and, given the high level of unique mutations, appears to detect ongoing mutational change.

Plasma cfDNA from blood samples taken at times of disease progression were most informative. B90-15 and B20-17 taken at point of significant disease progression were most useful whilst B14-16 and B91-16, taken at points of less significant disease progression were less so. Those samples taken when disease was well controlled, B44-15, B31-15, B143-16, had few detectable mutations and contributed little to the understanding of tumour genetics. This most likely occurred because ctDNA levels were higher at times of disease progression and therefore there was more tumour derived DNA to sequence.

WES was of insufficient depth to pick up the *ESR1* mutations previously identified by targeted NGS. The emergence of these two mutations showed resistance to the letrazole treatment given at the time. This shows the importance of high depth of sequencing in evaluating evidence for disease resistance in total cfDNA as well as breath of coverage as with WES. Depth of 1000-2000 is required to detect mutations present at <1% frequency.

The mutational signature profiles show that carboplatin chemotherapy, which has high efficiency in *BRCA* associated cancers that predominantly have the MS-3 profile, was given appropriately as the preceding total cfDNA sample (B14-16) had a high level of MS-3. SCNA

analysis of the WES data showed the patient was not *ERBB2* amplified in sample B90-15. This result predicted the non-response to TDM-1 therapy which is an anti-HER2 antibody-chemotherapy conjugate drug indicated for use in HER2 positive breast cancers only (Verma et al., 2012). Since sample B90-15 also had a strong MS-3 profile carboplatin could have been at this time, instead of TDM-1, 1 month earlier. Acting on such information, and thereby selecting the right drug at the right time, may have prevented clinical deterioration. Pacrintib may also have been active in this patient given the presence of a 1q21.3 amplification, and in the future patients with *MYC, SOX2 or TP53BP2* amplification, *PIK3CA* mutations or the other identified tier 2 mutations may benefit from yet to be approved therapies. These findings show the significant potential for total cfDNA WES analysis in the development of personalised medicine.

5 The genetics of gastroesophageal adenocarcinoma and the use of circulating cell free DNA for its detection and monitoring

5.1 Introduction

This chapter investigates the utility of ctDNA by SNV and SCNA analysis in 44 patients with GOA, including 13 metastatic patients, 27 patients with initially localised disease and 4 exceptional responders to chemotherapy.

5.2 Hypothesis, Aims and Objectives

5.2.1 Hypothesis and Aims

The aims of this chapter were to test the hypothesis that ctDNA can be detected in patients with GOA by molecular profiling. In addition, ctDNA was predicted to have utility in the monitoring of disease and the detection of minimal residual disease.

5.2.2 Objectives

5.2.2.1 Tumour tissue and total cfDNA analysis by NGS.

1. Design a custom NGS panel to detect common mutations in GOA tumours.
2. Compare paired tissue DNA / total cfDNA analysis by ultra deep sequencing using the customised NGS panel to detect ctDNA.
3. Determine changes in ctDNA over time during anticancer therapy, using a combination of
    i. Ampliseq total cfDNA analysis
    ii. Total cfDNA analysis by ddPCR for patient-tumour specific mutations

5.2.2.2 SCNA analysis of Tumour tissue and total cfDNA.

4. Use NanoString™ nCounter to detect SCNAs in the tumour tissue of patients with GOA undergoing curative treatment who lack a detectable SNV in their tumour tissue.
5. Compare paired tissue DNA / total cfDNA SCNA analysis by ddPCR and qPCR analysis to detect ctDNA.
6. Determine the utility of total cfDNA SCNA analysis to monitor patient therapy in the curative intent setting.

5.2.2.3 Patients with HER2 positive disease

7. Determine the utility of *ERBB2* SCNA analysis of tumour tissue and total cfDNA using ddPCR and qPCR, in patients with HER2 positive tumours, as assessed by IHC and FISH.

5.3 Patient recruitment and sample size

Patients with GOA were recruited from the gastrointestinal new patient clinic run by Professor Thomas. This included patients who presented with metastatic disease and patients with stage 2 and 3 tumours that were amenable to curative surgery. Patients known to be excellent responders to palliative chemotherapy, defined as survival greater than 5 years after diagnosis with no active disease, were also recruited.

The proportion of patients predicted to have detectable ctDNA in their blood samples was unknown. Formal power calculations were therefore not carried out as this was a proof of concept study. However, sample size was guided by previous studies.

In their study on Breast cancer Rothe et al. (2014) investigated 17 metastatic breast cancer patients to investigate the utility of ctDNA as an alternative to biopsy in breast cancer. Due to the limited degree of previous ctDNA studies in GOA, 17 was the minimum number of metastatic patients to be recruited.

As GOA has a higher relapse rate of approximately 60% at two years (Cunningham et al., 2006), 26 patients would need to be recruited in the neoadjuvant/adjuvant setting to get a 15 relapses for total cfDNA analysis in GOA cancers.

A total of 44 patients with GOA were recruited including; 27 patients with the initial aim of treatment with curative intent, of whom 23 went on to have curative surgery and 13 metastatic patients (

Figure 38). The 4 patients initially treated with neoadjuvant chemotherapy but who did not undergo surgery were followed up with the 13 metastatic patients, forming a palliative treatment cohort of 17 patients. 4 Patients that had had an exceptional response to chemotherapy were also recruited.

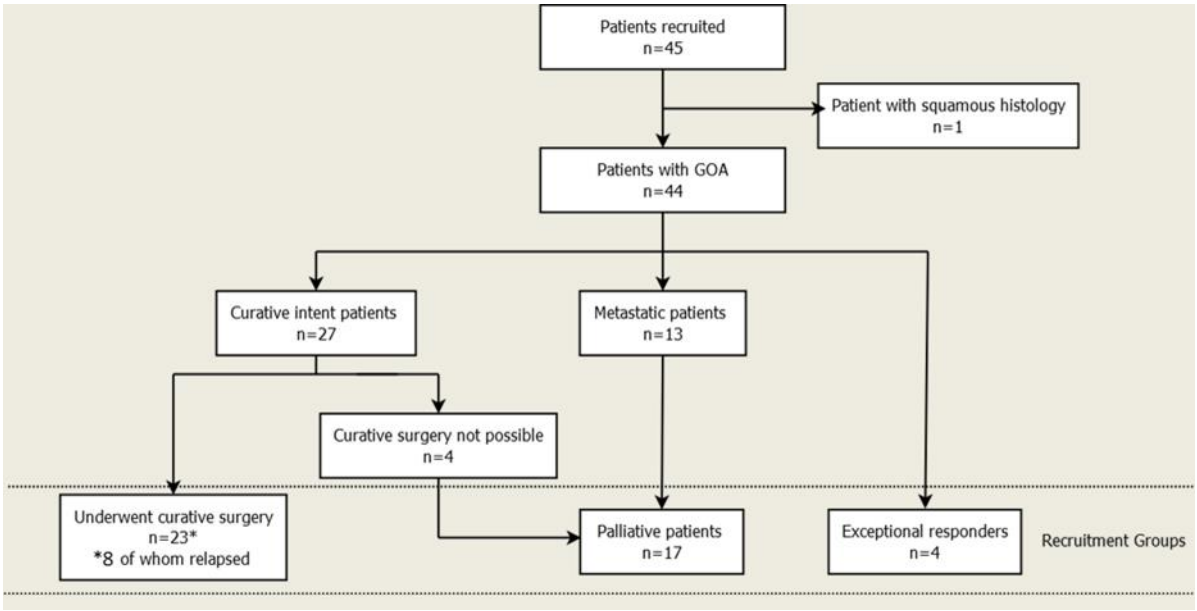


Figure 38 – **Consort diagram showing the recruitment of patients with gastroesophageal cancer.**

The clinicopathological details of the patient cohort are presented in Table 24.

	Palliative Patients		Curative Intent		Exceptional Responders	
	No.	%	No.	%	No.	%
Total	17	100	23	100	4	100
Histology						
Adenocarcinoma	15	88	20	87	4	100
Adeocarcinoma Diffuse Subtype	2	22	3	13	0	0
Type of GOA						
Gastric	6	35	4	18	1	25
GOJ type 1	4	24	6	26	1	25
GOJ type 2	3	18	3	13	0	0
GOJ type 3	2	11	4	17	0	0
GOJ type unclear	1	6	4	17	1	25
Oesophageal	1	6	2	9	1	25
Lymph node Status at Dx						
N>1	14	83	21	91	3	75
0	2	11	2	9	1	25
Unknown	1	6	0	0	0	0
Tumour size at Dx						
TX,0,1	0	0	1	4	1	20
T2	1	6	3	13	0	0
T3	13	77	17	75	3	60
T4	2	11	1	4	1	20
Unknown	1	6	1	4	0	0
Tumour Grade at diagnosis						
I	1	6	0	0	0	0
II	5	29	13	57	0	0
III	7	41	9	39	1	25
Unknown	4	24	1	4	3	75
Stage at initial diagnosis						
1	1	6	2	9	0	0
2	0	0	4	18	0	0
3	6	35	16	69	0	0
4	10	59	0	0	4	100
Not available	0	0	1	4	0	0
HER 2 Status						
Positive	3	18	0	0	0	0
Negative	11	64	0	0	1	25
Unknown	3	18	23	100	3	75
Test Pending	0	0	0	0	0	0
Site of metastases *	at diagnosis		at relapse		at diagnosis	
Lung	3	18	1			
Liver	4	24	1			
Peritoneal	4	24	1		1	25
Other	2	11	4		2	50
No metastases	7†	41			1	25

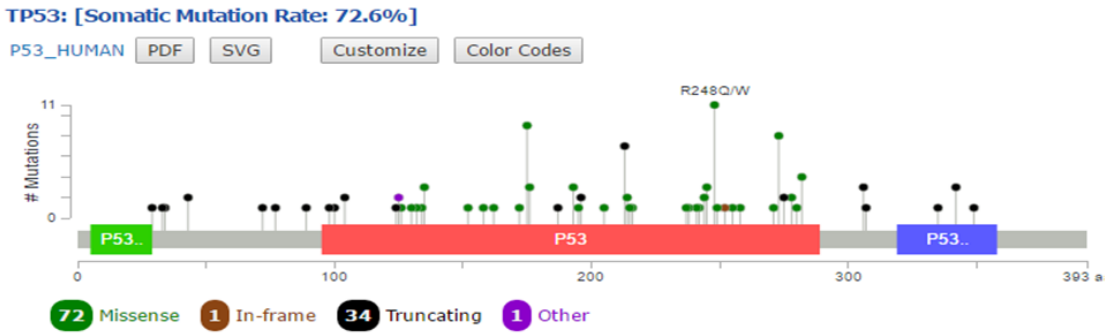
Table 24 – **Clinicopathological characteristics of 44 patients.** *Sites of metastases may be multiple and therefore the total count may exceed the number of patients. †Initially diagnosed with non-metastatic disease, which was inoperable or progressed to metastatic disease.

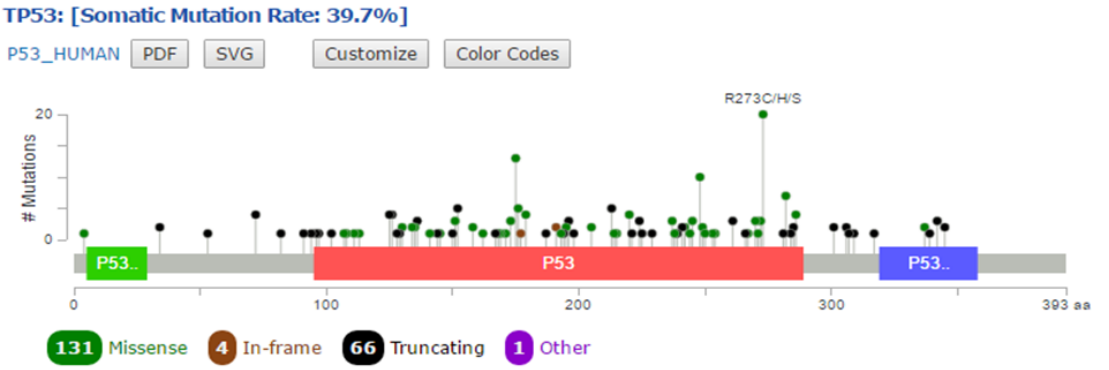
5.4 Design of Ampliseq targeted GOA panel

Using the WES and WGS data summarised in Figure 4 in the introduction (page 42), the 6 most commonly mutated genes in GOA were identified for inclusion in a custom NGS panel (*TP53, PIK3CA*, *ARID1A, RHOA, SMAD4* and *KRAS)*. Sites of recurrent SNVs and regions where SNVs clustered were identified using the website cBioportal (http://www.cbioportal.org). cBioportal included data from Cancer Genome Atlas Research (2014) and Dulak et al. (2013) discussed in the introduction.

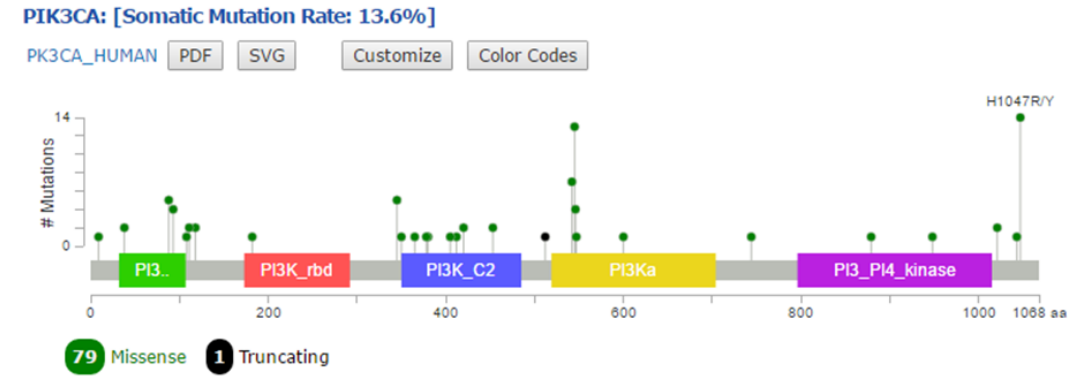
Mutations in *TP53* were identified evenly across the gene, with a few areas containing a higher frequency of mutations called hotspots (Figure 39a). Therefore, the Ampliseq panel was designed to cover the majority of the gene. Since the AcquiRes breast panel (section 2.13.1) already had significant coverage of *TP53,* 5 amplicons from that panel were included in the designed GOA panel (see Table 25).

a)





b)



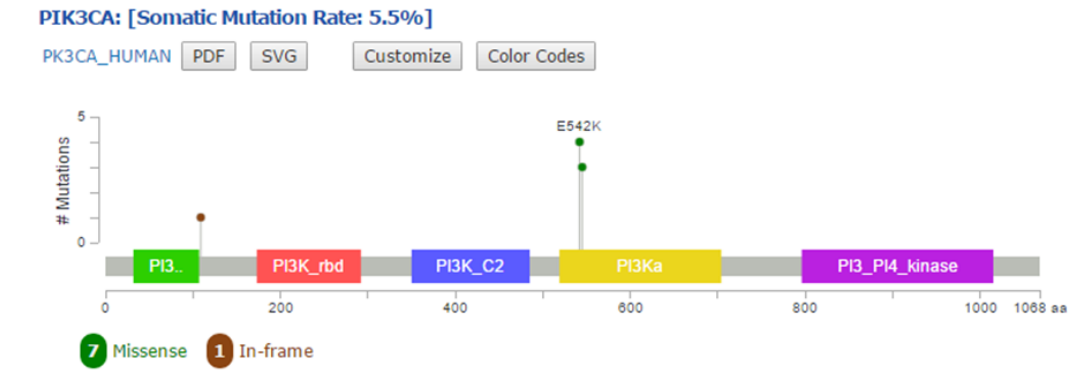


Figure 39 – **Distribution of mutations from cBioportal** a) TP53 with hotspots at p.R248Q and p.R273C b) PIK3CA with hotspot at p.E542K and P.H1047R. Top diagram in each section is data from Dulak et al. (2013), bottom diagram is data from Cancer genome atlas (2016).

In contrast mutations in *PIK3CA* were clustered at specific sites, and more common in patients with GOJ adenocarcinoma / OA (Figure 39b). This allowed focal areas of the *PIK3CA* gene to be targeted in 9 amplicons in the NGS panel.

The 4 other genes included in the panel were *ARID1A, RHOA, SMAD4* and *KRAS. KRAS* was included specifically because it has two hotspots that could be covered by three amplicons, and therefore could be justified in its inclusion despite having an overall low frequency of mutation. The coordinates for the included amplicons are shown (Table 25) and were designed through the Ion Ampliseq Designer (https://ampliseq.com). This used the DNA-hotspot design (single pool) for hg19, specifically for total cfDNA with amplicon size of 125-140bp. The algorithm behind how these amplicons are calculated is proprietary technology to Thermofisher who supply the programme. Following generation of the design, the 5 amplicons for TP53 highlighted below (Table 25) were edited into the panel following liason with the manufacturer.

	HG19 Chromosomal Co-ordinate		
Chromosome	Start Location	End location HG19	Amplicon
chr1	27056150	27056238	ARID1A_1
chr1	27087452	27087538	ARID1A_2
chr1	27088603	27088700	ARID1A_3
chr1	27089665	27089757	ARID1A_4
chr1	27097593	27097682	ARID1A_5
chr1	27099863	27099956	ARID1A_6
chr1	27100142	27100235	ARID1A_7
chr1	27100280	27100367	ARID1A_8
chr1	27101006	27101090	ARID1A_9
chr1	27105481	27105567	ARID1A_10
chr1	27105855	27105943	ARID1A_11
chr1	27106565	27106663	ARID1A_12
chr1	27106794	27106889	ARID1A_13
chr3	49405886	49405979	RHOA_1
chr3	49412866	49412952	RHOA_2
chr3	49412996	49413088	RHOA_3
chr3	178916704	178916783	PIK3CA_1
chr3	178916852	178916939	PIK3CA_2
chr3	178917409	178917491	PIK3CA_3
chr3	178921499	178921570	PIK3CA_4
chr3	178927933	178928015	PIK3CA_5
chr3	178928071	178928158	PIK3CA_6
chr3	178936029	178936108	PIK3CA_7
chr3	178951933	178952023	PIK3CA_8
chr3	178952070	178952153	PIK3CA_9
chr12	25378526	25378610	KRAS_1
chr12	25380263	25380351	KRAS_2
chr12	25398233	25398308	KRAS_3
chr17	7573994	7574075	TP53_1
chr17†	7577013†	7577142	TP53_2
chr17†	7577495†	7577632	TP53_3
chr17†	7578178†	7578307	TP53_4
chr17†	7578346†	7578484	TP53_5
chr17†	7578514†	7578650	TP53_6
chr17	7579287	7579381	TP53_7
chr18	48591749	48591837	SMAD4_1
chr18	48591878	48591964	SMAD4_2
chr18	48593402	48593481	SMAD4_3
chr18	48604621	48604710	SMAD4_4

Table 25 – **Amplicons with Hg19 genomic co-ordinates comprising the designed ampliseq GOA panel.** †Denotes amplicons taken from the AquiRes panel.

5.5 Results: Optimisation of GOA Panel

5.5.1 Comparing 2 NGS amplicon panels; the AcquiRes and GOA panel

Since the majority of the *TP53* amplicons were taken from the AcquiRes panel (See Table 25 and Table 8), an initial sequencing run of 8 GOA tumour DNA samples was carried out using the AcquiRes panel to:

1. Confirm that *TP53* is commonly mutated in the GOA.
2. Allow a test run of the same tumour samples to ensure concordance between the GOA and AcquiRes panel. Thereby acting as a control to ensure the GOA panel is working correctly.

TP53 mutations were detected in 6 of the 8 tumour samples using the AcquiRes panel, and one sample (GT02.1) also had an additional *PIK3CA* mutation (Table 26). All samples had good depth (mean reads per amplicon) and uniformity (a measure of variation in read-depth across each base / amplicon). Both FFPE and fresh frozen samples were sequenced and results showed good read depth and quality.

Run summary matrix							SNV specific matrix					
Sample	Patient Number	Type of Sample	Mapped Reads	On Target	Mean Depth	Uniformity	Gene	Mutation	Coverage	Mutant Call	Freq %	Quality
GT02.1	CRB055	FFPE	181609	95.4	5373	96.2	TP53	R196X	5816	4486	77.6	56498
							PIK3CA	I338N	5051	1946	39.0	17192
B35/16.3	CRB058	FF	130679	96.6	3832	100.0	None	-				
B36/16.4	CRB059	FF	271633	94.2	8171	100.0	TP53	C277F	6000	2611	56.8	36809
GT12.1	CRB062	FFPE	217092	98.1	6425	96.2	TP53	R280K	2623	731	27.8	5202
GT01.1	CRB063	FFPE	31987	96.1	943.7	96.0	TP53	R282W	392	130	33.0	1521
GT07.1	CRB066	FF	756310	95.0	21829	99.9	None	-				
GT19.1	CRB070	FFPE	161945	95.4	4777	100.0	TP53	R213X	4350	257	6.0	296
GT40.1	CRB081	FFPE	377865	94.7	10639	96.2	TP53	R306X	4704	638	13.6	2506

Table 26 - **Sequencing metrics and mutations detected from 8 tumour DNA samples using the AcquiRes ampliseq panel** . FF = Fresh Frozen sample. Run summary matrix (left) shows overview statistic for each run. SNV specific matrix (right), shows statistics for the specific base causing the mutation.

The same 8 tumour DNA samples were then sequenced with the GOA panel. The same *TP53* and *PIK3CA* mutations were detected at a similar frequency, within a 2-5% margin (Table 27). One additional *KRAS* mutation was also detected (sample B35/16.3), not covered by the AcquiRes panel. The sequencing metrics were also compared and this showed an overall lower quality than with the AcquiRes panel, based on poorer overall mean depth, uniformity and coverage (Table 27).

Run summary matrix						SNV specific matrix					
Sample	Patient #	Mapped Reads	On Target %	Mean Depth	Uniformity %	Gene	Mutation	Coverage	Mutant Call	Freq %	Quality
GT02.1	CRB055	102640	98.95	2440	80.88	TP53	R196X	333	239	71.9	2985
						PIK3CA	I338N	1930	718	37.6	4503
B35/16.3	CRB058	122182	99.44	2957	93.87	KRAS	G12D	1617	605	37.4	5168
B36/16.4	CRB059	83411	99.23	2023	78.82	TP53	C277F	557	346	62.2	3912
GT12.1	CRB062	177689	99.70	4233	62.62	TP53	R280K	328	107	32.8	820
GT01.1	CRB063	50028	98.96	1241	95.33	TP53	R282W	572	216	38.1	1878
GT07.1	CRB066	601078	98.04	14457	97.62	None	-				
GT19.1	CRB070	336890	99.16	8554	90.55	TP53	R213X	5322	279	4.9	527.2
GT40.1	CRB081	493468	97.95	11744	97.62	TP53	R306X	4210	481	11.4	1487

Table 27 – **Sequencing metrics and mutations detected in the first sequencing run using the GOA ampliseq panel**. Run summary matrix (left) shows overview statistic for each run. SNV specific matrix (right), shows statistics for the specific base causing the mutation.

5.5.2 Optimisation of the GOA Panel; buffy coat and FFPE DNA samples

Buffy coat (lymphocyte) germline DNA was used for optimisation initially as this is higher molecular weight DNA than total cfDNA or FFPE. The patients with samples with a mutation in section 5.5.1 had their corresponding buffy coat DNA samples sequenced. This also served as a check for any germline variants.

Run summary matrix						SNV specific matrix					
Buffy Coat Sample	Patient #	Mapped Reads	On Target %	Mean Depth	Uniformity %	Gene	Mutation	Coverage	Mutant Call	Freq %	Quality
B110/15	CRB055	539013	99.03	12949	86.82	TP53	R196X	12541	0	0	0
						PIK3CA	I338N	3024	5	0	0
B118/15	CRB058	378752	98.94	9054	85.55	KRAS	G12D	8163	4	0	0
B119/15	CRB059	180346	98.54	4167	68.51	TP53	C277F	227	0	0	0
B124/15	CRB062	31292	94.32	662.2	85.11	TP53	R280K	121	0	0	0
B125/15	CRB063	88646	97.23	1967	64.39	TP53	R282W	23	1	4	0
B015/16	CRB070	227727	98.89	5271	66.13	TP53	R213X	153	0	0	0
B033/16	CRB081	185591	98.43	4252	63.26	TP53	R306X	158	0	0	0

Table 28 – **Sequencing metrics of buffy coast (germline) samples using the GOA panel with 25 multiplex PCR cycles.** Run summary matrix (left) shows overview statistic for each run. Coverage (on right) of the SNV locations identified in Table 27.

Results confirmed the mutations detected in the tumours were not present in the matched buffy coat samples (Table 28). However, results showed overall lower uniformity than for the matched tumour samples. As uniformity is a measure of variation in read-depth across each base, a low uniformity suggested some regions or amplicons were sequenced better than others. In keeping with this, amplicon coverage in amplicons 2-6 of *TP53* was much lower when uniformity was low (Table 29), and coverage for sample B125/15 at the site of the tumour specific mutation was just 23 reads (Table 28).

Sample →	B110/15	B118/15	B119/15	B124/15	B125/15	B15/16	B33/16
Amplicon ↓	Number of reads						
TP53 -1	24853	17665	12615	1496	6120	15420	13496
TP53-2	4297	2975	300	173	109	312	195
TP53-3	5829	3360	249	212	104	180	162
TP53-4	3188	1508	183	192	119	246	183
TP53-5	6343	3762	260	268	116	257	206
TP53-6	1710	1299	600	294	419	479	615
TP53-7	19455	12561	3229	474	670	2890	2000

Table 29 – ***TP53*** **amplicon coverage of buffy coat sequencing shown in**

Table 28**.** Highlighted amplicons TP53-2 to TP53-6 show low number of overall reads.

The sequenced protocol used 25 multiplex cycles for buffy coat whereas the optimised number was 20 for the AcquiRes panel (Table 6, page 68). Therefore, these samples were

re-sequenced with the GOA panel at 20 multiplex cycles for comparison. The results showed an improved uniformity and coverage of the SNV sites of interest (Table 30), and subsequent sequencing for buffy coat was therefore adjusted to 20 cycles. Fresh frozen tissue DNA, which is also of high molecular weight and quality was also run at 20 cycles.

Run summary matrix						SNV specific matrix					
Sample	Patient #	Mapped Reads	On Target %	Mean Depth	Uniformity %	Gene	Mutation	Coverage	Mutant Call	Freq %	Quality
B110/15	CRB055	355998	96.5	8708	97.62	TP53	R196X	2193	0	0	0
						PIK3CA	I338N	9975	1	0	0
B118/15	CRB058	216681	97.0	5316	97.62	KRAS	G12D	4522	4	0	0
B119/15	CRB059	368704	97.0	8922	97.62	TP53	C277F	5326	0	0	0
B124/15	CRB062	350596	97.8	8517	97.62	TP53	R280K	4054	1	0	0
B125/15	CRB063	380976	98.8	9413	97.62	TP53	R282W	5701	4	0	0
B015/16	CRB070	306464	96.5	7377	97.62	TP53	R213X	6254	2	0	0
B033/16	CRB081	374220	98.1	9204	97.62	TP53	R306X	4629	0	0	0

Table 30 - **Sequencing metrics of buffy coast (germline) samples using the GOA panel, with 20 multiplex PCR multiplex**. Including (on right) coverage of the mutation locations identified in Table 27. Repeated samples from

Table 28 with 20 cycles of multiplex PCR.

Given the finding that the buffy coat samples were sensitive to the number of cycles in the multiplex PCR, 3 FFPE samples were sequenced at 22 and 25 multiplex cycles for comparison of sequencing metrics.

Sample	Patient #	Multiplex Cycles	Mapped Reads	On Target	Mean Depth	Uniformity %
GT34.2-11	CRB088	22	212501	97.67%	5270	95.03%
GT34.2-12	CRB088	25	70432	99.18%	1694	80.89%
GT39.1-13	CRB109	22	500063	98.56%	11859	96.89%
GT39.1-14	CRB109	25	258708	97.95%	5763	77.29%
GT46.2-15	CRB110	22	342826	98.79%	8236	97.57%
GT46.2-16	CRB110	25	82112	98.39%	1941	73.54%

Table 31 – **Run summary matrix for FFPE samples repeated at 22 and 25 multiplex cycles.**

The results showed increased mean depth and uniformity when 22 multiplex cycles were used (Table 31). As an example, the *KRAS, TP53* and *SMAD4* amplicons showed better coverage at 22 multiplex cycles (Figure 40). Of note the amplicons taken from the AcquiRes panel (TP53-2 to TP53-6) generally had less coverage, likely due to the fact that these amplicons were on average 30bp longer than the other amplicons. Overall, reducing the number of multiplex cycles improved the coverage of all amplicons. Therefore, subsequent sequencing of FFPE samples used 22 multiplex cycles.

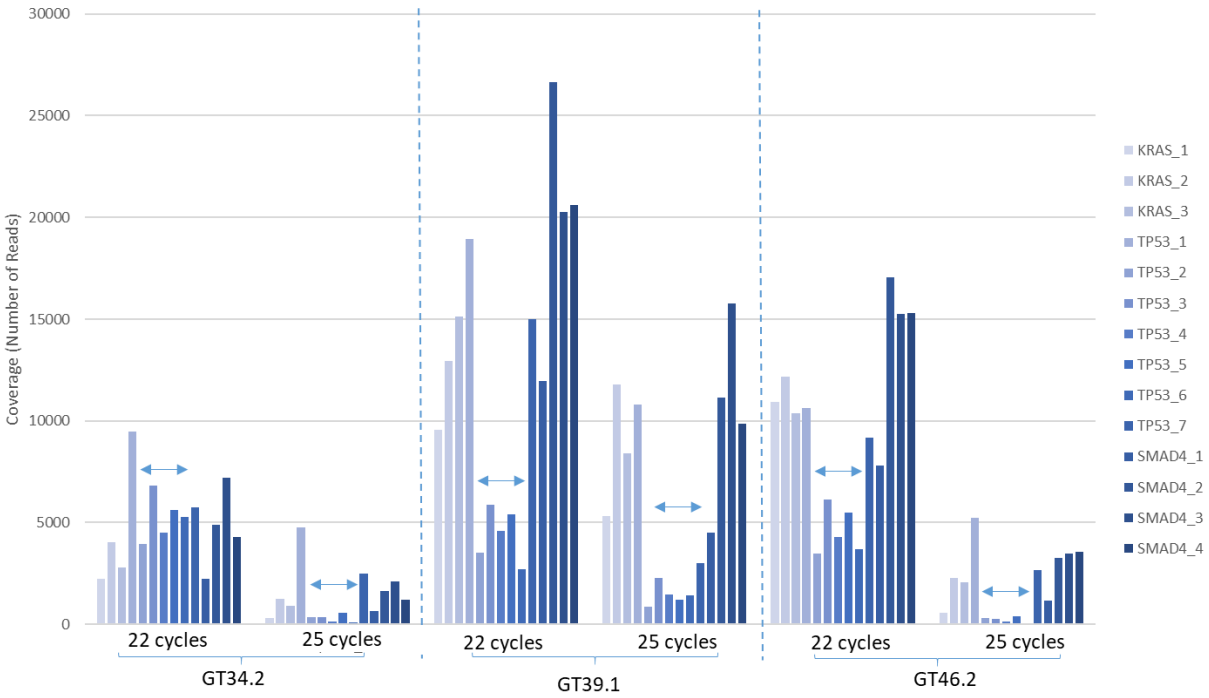


Figure 40 - **Bar chart showing amplicon coverage of 3 FFPE samples following 22 and 25 cycles of multiplex PCR with the GOA panel**. The amplicons covering *KRAS, TP53* and *SMAD4* are shown. Arrows indicate the presence of TP53 amplicons taken from the AcquiRes panel.

5.5.3 FFPE sequencing

Using the optimised protocol, a total of 20ng of DNA was sequenced with the GOA panel from the primary tumour (36 FFPE, 7 Fresh Frozen, 1 tumour DNA unavailable) of 44 patients with GOA. 4 patients had insufficient DNA from their tumour samples or the quality of the DNA was too low to proceed (Figure 41, Table 32). Overall results were generated for 40 patients of whom 33 had 1 or more detectable mutations at >5 % VAF.

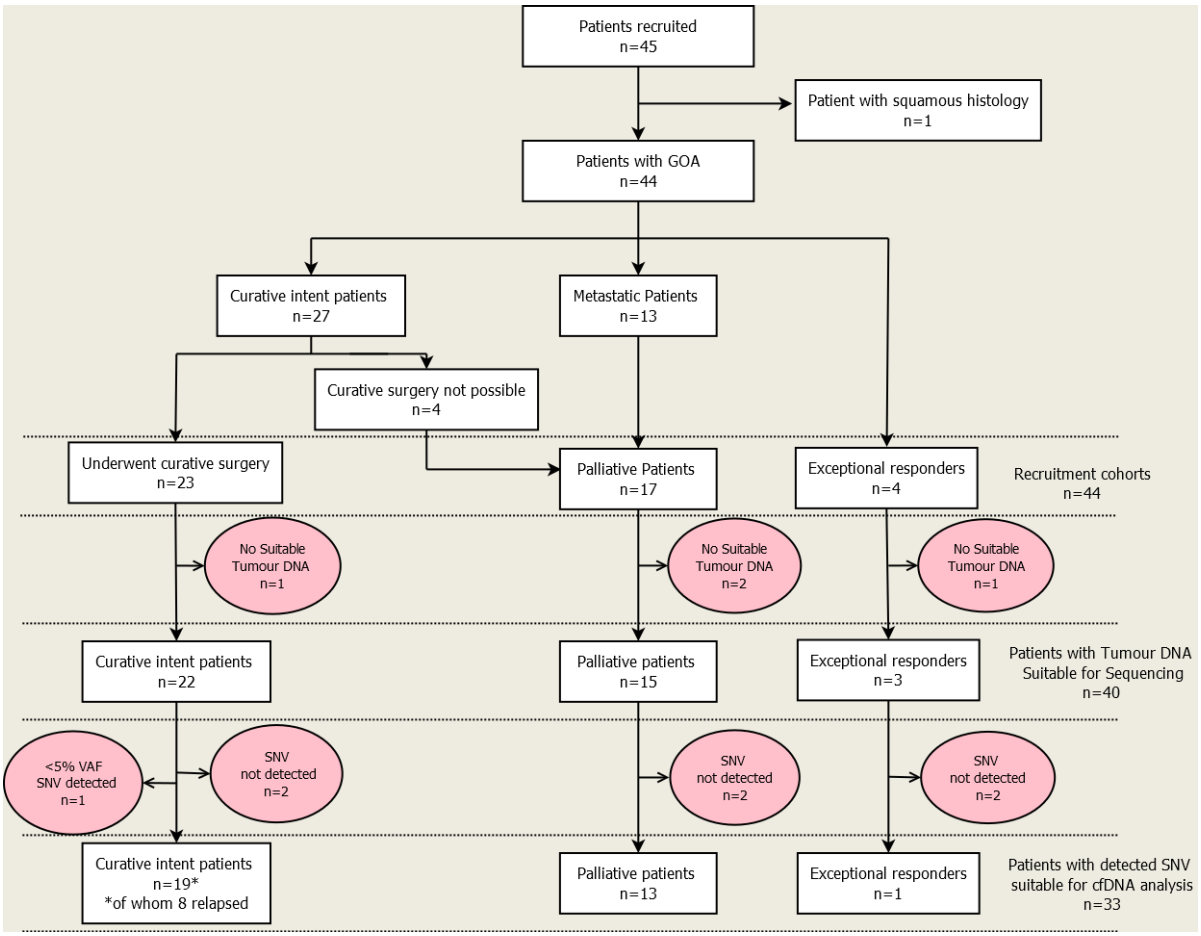


Figure 41 – **Consort Diagram showing patient work flow and number of tumour samples sequenced.**

	Patient	Type	Source DNA	Mean Depth	Gene	Mutation	Coverage	Mutant Call	VAF (%)	Quality	Confimed by ddPCR
1	CRB55	C	FFPE	2440	TP53	R196X	333	239	71.9	2985	Yes
					PIK3CA	I338N	1930	718	37.2	4503	No
2	CRB58	C	FF	2957	KRAS	G12D	1617	605	37.4	5168	Yes
3	CRB59	C	FF	2023	TP53	C277F	557	346	62.2	3912	Yes
4	CRB61	C	FF	12314	TP53	R175H	7743	2282	29.5	14269	Yes
5	CRB62	P	FFPE	4233	TP53	R280K	328	107	32.6	820	Yes
6	CRB63	P	FFPE	1241	TP53	R282W	572	216	37.8	1878	Yes
7	CRB64	P	FFPE	8984	TP53	R248Q	1960	250	12.8	1442	Yes
8	CRB66	Ex	FFPE	14457	No mutation						
9	CRB67	P	FFPE	3012	TP53	R175H	661	358	54.2	4552	Yes
10	CRB69	P	FFPE	5842	TP53	E271X	4397	1999	45.5	10265	Yes
11	CRB70	C	FFPE	8554	TP53	R213X	5322	279	5.2	527	Yes
12	CRB71	C	FFPE	4052	TP53	R248Q	2431	857	35.3	7203	Yes
13	CRB74	P	FFPE	6072	TP53	S241A	4168	395	9.4	566	Yes
14	CRB75	P	FFPE	3488	TP53	R213X	320	94	29.4	863	Yes
15	CRB78	C	FFPE	4938	No mutation						
16	CRB80	C	FFPE	15949	TP53	R175H	3404	218	6.4	464	Yes
17	CRB81	Ex	FFPE	11744	TP53	R306X	4210	481	11.4	1487	Yes
18	CRB82	C	FFPE	9621	TP53	R306X	1927	378	19.6	4316	Yes
19	CRB83	P	FFPE	2863	TP53	R249G	3172	947	29.9	7951	No
20	CRB84	P	FFPE	2423	TP53	P151S	95	53	55.8	689	Yes
21	CRB85	P	FFPE	6556	SMAD4	R361S	7017	3077	43.9	25612	No
22	CRB86	P	FFPE	9402	No mutation						
23	CRB88	C	FFPE	5270	TP53	E286K	3816	1392	36.4	12263	Yes
24	CRB89	P	FFPE	6220	TP53	F113C	10002	2521	25.2	11242	Yes
25	CRB94	C	FF	11017	No mutation						
26	CRB95	C	FF	3807	TP53	P151R	617	328	53.2	4577	Yes
					ARID1A	Y1319X	4517	1990	44.2	17392	No
27	CRB99	P	N/A		Tumour sample unavailable						
28	CRB102	C	FF	4302	ARID1A	Q386X	13366	441	3.3	266	No
29	CRB105	C	FF	7213	TP53	C135F	2585	1448	56.0	14992	Yes
30	CRB106	Ex	FFPE	4997	No mutation						
31	CRB107	C	FF	8532	TP53	Q192X	3009	1848	61.5	23593	No
32	CRB108	P	FFPE		Tumour DNA of too low quality						
33	CRB109	P	FFPE	11859	No mutation						
34	CRB110	C	FFPE	8236	TP53	C176F	5377	1500	27.8	12614	Yes
35	CRB113	Ex	FFPE		Tumour DNA of too low quality						
36	CRB114	C	FFPE		Tumour DNA of too low quality						
37	CRB115	C	FFPE	4155	TP53	H179R	2497	153	6.0	665	Yes
					RHOA	A44V	3803	342	9.0	943	No
38	CRB116	C	FF	6903	TP53	G266V	2550	1483	58.2	14474	Yes
39	CRB117	C	FFPE	5009	TP53	R213X	2212	460	20.9	4330	Yes
40	CRB118	P	FFPE	7777	TP53	R213X	3536	864	24.4	5106	Yes
41	CRB121	P	FFPE	4537	TP53	V122Dfs26X	1957	879	44.9	9152	Yes
42	CRB137	C	FFPE	13071	TP53	T125T	24704	21911	88.7	81339	Yes
43	CRB144	C	FFPE	2205	TP53	S215R	365	210	57.5	2438	Yes
44	CRB145	C	FFPE	13573	TP53	E286K	1366	465	34.0	4155	Yes
					PIK3CA	E545K	14058	901	6.4	896	No

Table 32 – **Summary matrix of all 44 sequenced tumours.** FF= Fresh Frozen Tissue. C=Curative Intent, P=palliative, Ex= exceptional responder

In total 31 of the 40 patients (77.5%) with sequencing results had a *TP53* mutation (Figure 42), which aligns with the 70-80% frequency from predicted other WES and WGS studies (Section 1.3.3, Figure 4, page 42). This suggests that the GOA panel was picking up the majority of *TP53* mutations.

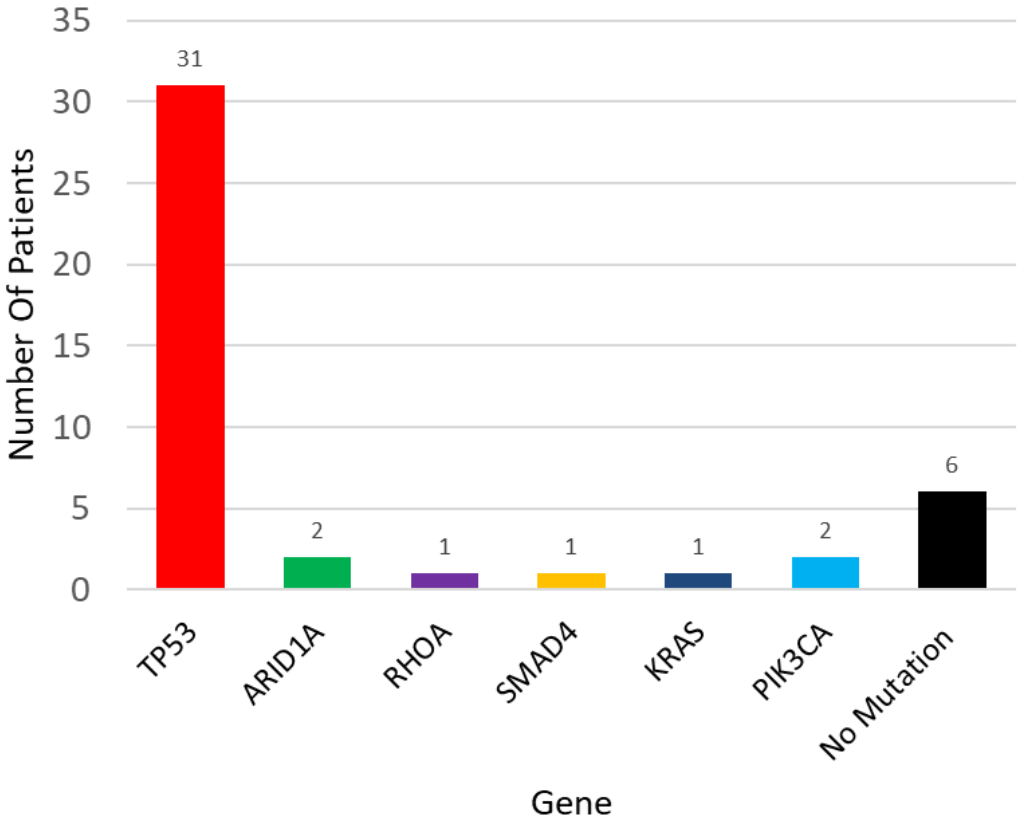


Figure 42 – **Bar chart showing the cumulative number of patients for each gene that have a mutation.** Note some patients have > 1 mutation and are therefore counted twice.

Considering mutation frequency all the mutations were present at >5% VAF, apart from in one tumour with an *ARID1A* p.Q386X mutation present at 3.3% VAF. *TP53* mutations were present with a VAF of 5.2-88.7% and a median of 32.8% (Figure 43).

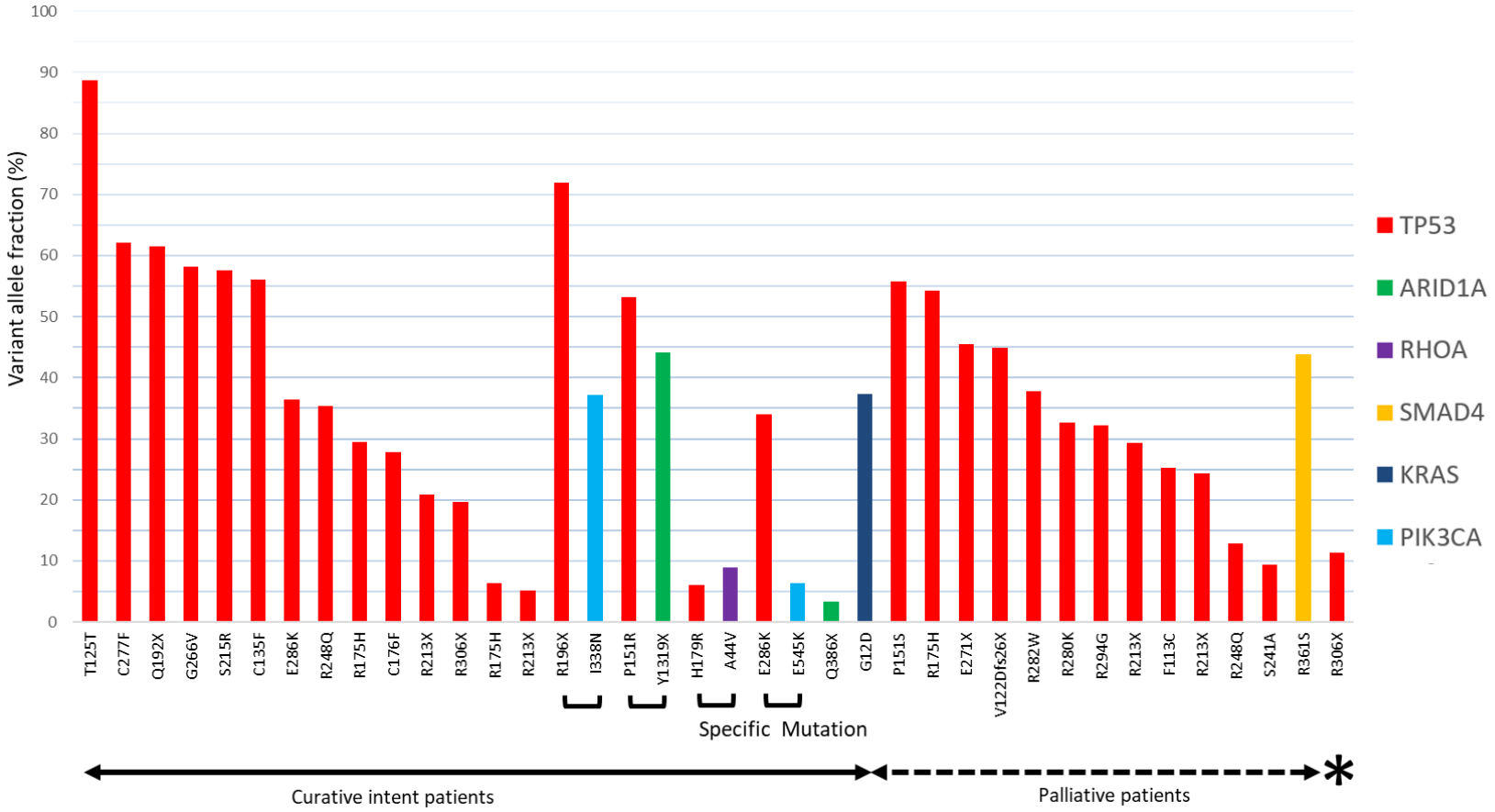


Figure 43 – **Frequency of specific mutation across all sequenced tumours.** Solid bar indicates two detectable mutations in a single tumour tissue DNA sample. *Exceptional responder.

5.5.4 Optimisation of total cfDNA amplification

Based on results from FFPE and FF tissue DNA, the panel was also optimised for cfDNA with the GOA panel. The first test was with 28 cycles of multiplex (as used for the AcquiRes panel) using 20ng of input DNA. This resulted in libraries with high concentration of 1.6-2.7ng/µl which can indicate over amplification. A comparison of sequencing at 28 and 26 multiplex cycles was therefore designed again using the same amount of input DNA. It was not possible to repeat one sample (CRB62 B06/17), as there was insufficient total cfDNA for experimentation, for that reason a different blood sample B41/16 was selected. Overall 26 cycles appeared to improve depth and uniformity compared to 28 cycles (Table 33).

In terms of optimisation, previous literature suggested the use of as few cycles of multiplex PCR as possible, as multiple cycles of PCR introduce amplification artefacts, amplification amplicon bias and duplicative reads (Kinde et al., 2011, Stasik et al., 2018). Amplification artefacts occur due to DNA polymerase errors, primer miss-binding and dimer formation. Amplicon bias may occur because some amplicons are amplified preferentially to others (Dohm et al., 2008). In addition, the ratio of the original mutant to wild-type DNA may be lost during amplification as some of the original DNA molecules may amplify preferentially to others (Peng et al., 2015) . Therefore, 26 multiplex cycles was decided on as the trade-off between having sufficient cycles to have enough product to sequence whilst keeping the error rate of sequencing to a minimum.

Sample	Patient #	Multiplex Cycles	Library prep Ng/ul	Mapped Reads	On Target %	Mean Depth	Uniformity %
B49/16	CRB87	28	1.9	107130	98.28	2656	58.29
B49/16	CRB87	26	1.17	183472	99.04	4489	86.60
B65/16	CRB109	28	2.71	121474	99.06	3053	85.33
B65/16	CRB109	26	0.812	126293	98.84	3030	71.99
B06/17	CRB62	28	1.66	19689	98.11	481.5	49.75
B41/16	CRB85	26	0.653	264049	99.52	6412	83.07

Table 33 – **26 versus 28 cycles of multiplex PCR for total cfDNA sequencing.** Using 20ng of input DNA. No variants were detected in these samples.

5.5.4.1 Analysis of total cfDNA samples using the GOA panel

Based on results from the study of palliative breast patients in chapter 3.3 it was decided to select those patients who had significant disease on CT for total cfDNA profiling with the GOA panel. Patient CRB62 was the first patient to be recruited at the metastatic stage and fitted the criteria. A *TP53* p.R280K mutation at 32.8% VAF was detected in the tumour DNA. 5 serial bloods taken over 60 weeks (421 days) were sequenced using 20ng of input DNA and 26 multiplex cycles. Only one sample (B06/17) showed evidence of ctDNA (*TP53* p.R280K 11.4% VAF) and was taken at a time of disease progression just before death (

Table 34). This confirmed the GOA panel could detect ctDNA. However, the variability and poor coverage from total cfDNA suggested uncertainty about whether the other blood samples, were true negatives.

Run summary matrix						R280K specific Matrix			
Blood ID	Blood #	Mapped Reads	On Target %	Mean Depth	Uniformity %	Coverage	Mutant Call R280K	Freq %	Quality
B124/15	1	349259	12.76	966.7	69.20%	19	0	0	NoCall
B39/16	2	81076	99.20	1913	62.57%	110	0	0	NoCall
B100/16	3	55310	99.56	1308	59.92%	40	0	0	NoCall
B148/16	4	89554	98.80	2112	60.19%	200	0	0	NoCall
B06/17	5	19689	98.11	481.5	49.75%	501	57	11.4	316

Table 34 – **Overview of ampliseq GOA panel sequencing of CRB62 total cfDNA samples.** 20ng input DNA and 26 multiplex cycles.

Therefore, each sample was also analysed with a commercially available *TP53* p.R280K ddPCR assay. The results showed samples B124/15, B100/16, B148/16 & B06/17 (bloods 1, 3,4,5) were positive for ctDNA determined by the presence of positive *TP53* p.R280K droplets. Sample B06/17 had the highest VAF (8.6%) in keeping with the results from sequencing (Figure 44). Results from the other samples had a VAF <2%, which suggested these samples were below the sensitivity of the GOA ampliseq panel. As controls, 2 tumour samples and buffy coat were also analysed. These results confirmed results from sequencing, with tumour samples being positive for *TP53* p.R280K and buffy coat samples being negative (Figure 44).

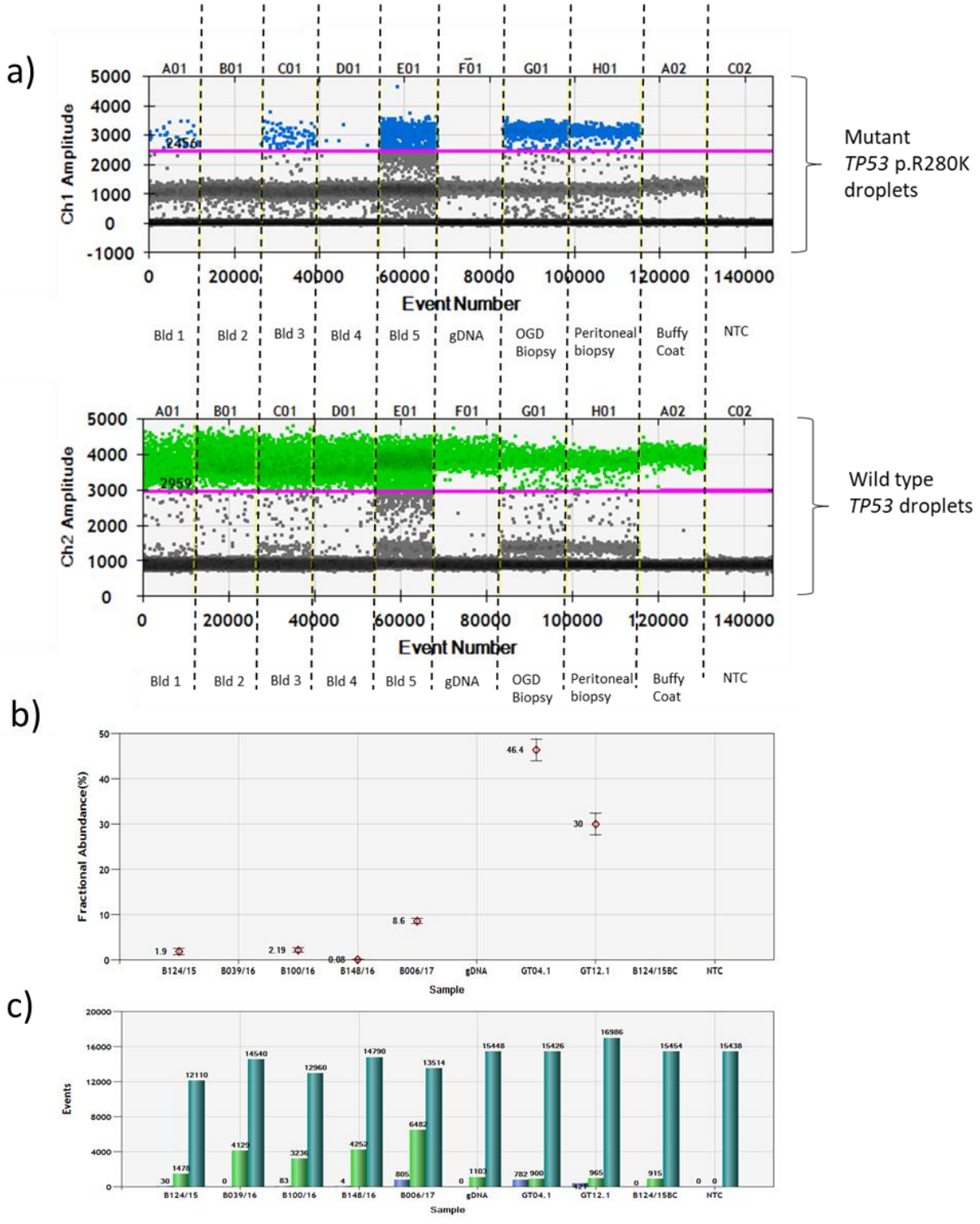


Figure 44 - **DdPCR analysis of total cfDNA, Tumour DNA and buffy coat DNA from patient CRB62**. a) Showing the raw data for the *TP53* p.R280K mutation, bloods 1-5, linked to Blood ID's as per Table 34. b) Fractional abundance = VAF and c) the total number of events (turquoise bar), and positive droplets for mutant (blue bar) and wildtype (green bar) DNA. GT04.1 = Tumour DNA from

OGD biopsy, GT12.1 = Tumour DNA from peritoneal biopsy, B124/15BC = Buffy coat DNA from initial blood sample B124/15, gDNA = genomic DNA, NTC = non-template control

As confirmation, NGS and ddPCR were also applied to samples from a second patient, CRB55, who was treated with curative intent but unfortunately relapsed at the time of the third blood test. The tumour DNA had a high frequency *TP53* mutation p.R196X (71.9% VAF) and 4 serial total cfDNA samples were analysed (Table 35). 3 of the four samples were positive by NGS and correlated with treatment (see section 5.6), and again the limit of detection was around 2% VAF. The results correlated well with the ddPCR results for the specific *TP53* p.196X mutation, apart from for sample B100/15 which had a lower VAF by ddPCR (Table 36). Of note sample B100/15 was run twice by sequencing in order to detect the p.R196X mutation, indicating that samples below a VAF of 2% are not reliably detected by the GOA panel.

Run summary matrix						R196X specific Matrix			
Blood ID	Blood #	Mapped Reads	On Target %	Mean Depth	Uniformity %	Coverage	Mutant Call R196X	Freq %	Quality
B100/15	1	293682	90.11%	5927	82.18%	1373	32	2.3	50
B16/16	2	311519	99.71%	7545	72.60%	774	0	0	NoCall
B77/16	3	212042	98.27%	4997	65.00%	252	46	18.2	386
B150/16	4	76191	98.75%	1818	63.76%	94	40	43.0	460

Table 35 - **Overview of ampliseq GOA panel sequencing of CRB55 total cfDNA samples.**

TP53 p R196X ddPCR					
Blood ID	Blood #	Total droplet	WT Droplets	Mutant Droplets	VAF %
B100/15	1	11351	1077	7	0.61
B16/16	2	14562	2363	1	0
B77/16	3	12253	628	138	17.8
B150/16	4	11658	1643	2127	43.1

Table 36 – **TP53 p R196X ddPCR for CRB55 total cfDNA samples**

To further investigate the limit of detection, a selection of total cfDNA samples were run at time-points of disease progression or just before death when it was predicted ctDNA was

more likely to be present. The results (Table 37) show that the lowest called mutation had a VAF of 2.5%. Although a 1.1% VAF mutation was present on review of the data for sample B97/16 this was not called by the variant caller program. This led me to infer that the limit of detection for this panel was around 2% when analysing total cfDNA samples.

Run summary matrix						SNV specific matrix					
Sample	Patient #	Mapped Reads	On Target %	Mean Depth	Uniformity %	Gene	Mutation	Coverage	Mutant Call	Freq %	Quality
B49/16	CRB087	183472	99.04%	4489	86.60%	TP53	C238R	2167	1	0	No call
B38/16	CRB083	39078	99.31%	1207	56.04%	TP53	R249G	89	7	7.9	49
B41/16	CRB085	264049	99.52%	6412	83.07%	SMAD4	R361S	6193	152	2.5	7.2
B97/16	CRB107	187700	92.24%	4067	80.64%	TP53	Q192X	947	11	1.1	No call
B157/16	CRB107	136730	99.14%	3889	56.16%	TP53	Q192X	289	69	23.9	355
B53/16	CRB089	661	0.7186	8.654	0.6648	TP53	F113C	4	1	N/A	No call
B111/16	CRB089	8228	0.9499	174.8	0.5406	TP53	F113C	18	1	5	No call
B130/16	CRB089	92090	96.91%	2060	59.86%	TP53	F113C	635	197	29.4	1457

Table 37 – **'High likelihood' total cfDNA samples sequenced using the ampliseq GOA panel.**

5.5.4.2 Conclusions from total cfDNA sequencing and plan for further total cfDNA analysis

Review of the literature confirms a 2% VAF for the lower limit of detection using ampliseq and the PGM™ Iontorrent technology (Lih et al., 2017) with a lower limit of 1% used for known hotspot mutations (Frenel et al., 2015). Given the results from the ddPCR, it was apparent that levels of ctDNA <1% are present. As the majority of patients had only a single mutation following tumour NGS, analysis of other mutations in total cfDNA was restricted to ddPCR, in order to achieve a higher sensitivity of ctDNA detection. The results of these experiments are summarised in the next section.

5.6 Results: SNV Analysis in ctDNA using NGS and ddPCR

Figure 45 below summarises the patient workflow and cfDNA analyses. 33 patients had at least one detectable SNV >5% VAF in tumour DNA. Of these, 16 patients had the same SNV(s) detected indicating the presence of ctDNA and 17 patients were negative for ctDNA.

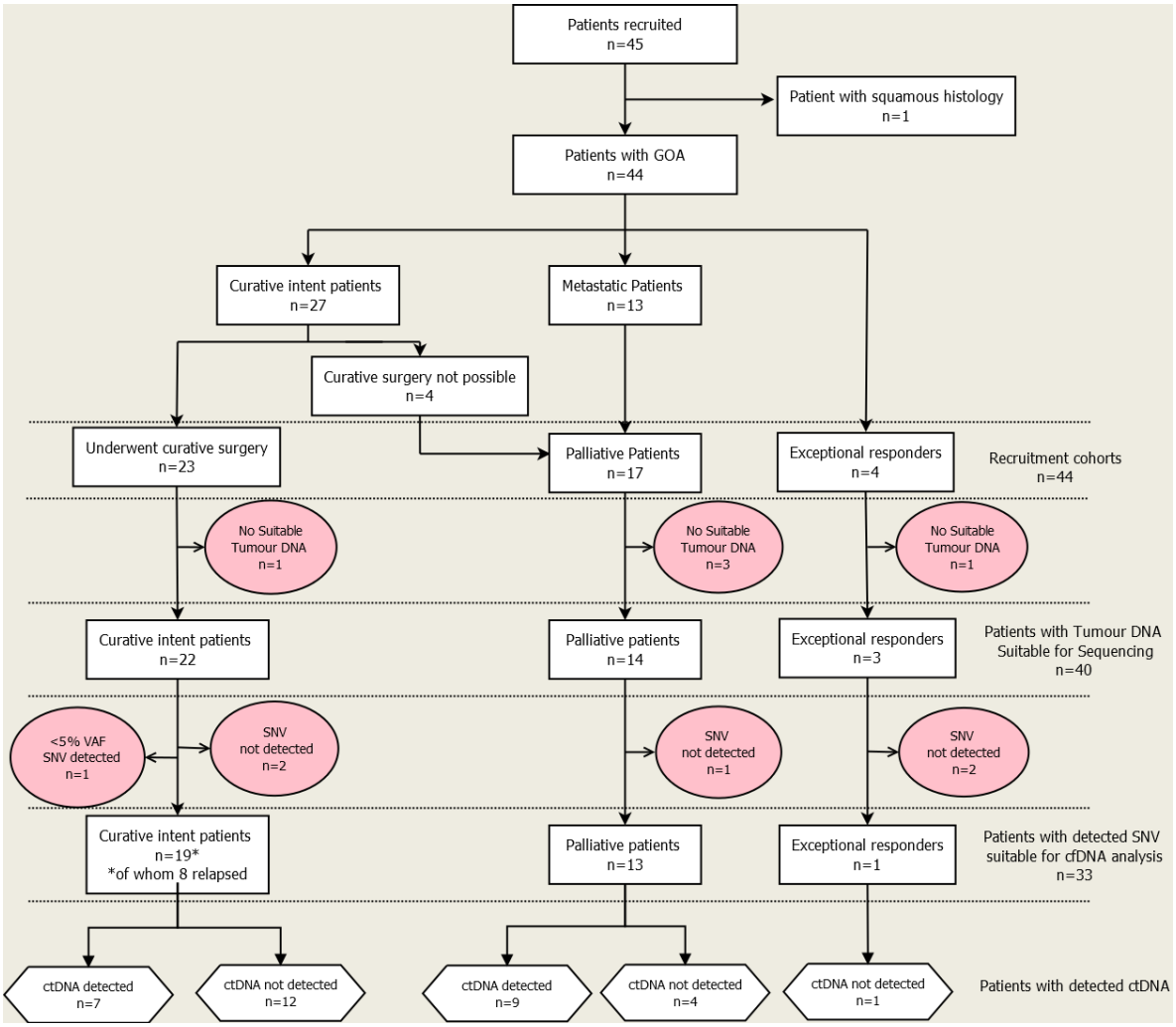


Figure 45 – **Consort diagram summarising patient recruitment and tumour DNA / total cfDNA analysis workflow.**

5.6.1 GOA palliative patients

Thirteen of 17 (76.4%) palliative patients had a mutation in their tumour (VAF>5%) and 9 of these 13 (69.2%) had the same mutation detected in total cfDNA, indicating the presence of ctDNA (Table 38). In those 9 patients the ctDNA levels (as determined by copies /ml) tracked disease response to treatment and progression of disease. The 4 most complete patient timelines for disease monitoring are discussed below. The remaining 5 timelines for cfDNA positive patients can be found in appendix 7.6.

Tumour samples were sequenced using the GOA ampliseq panel, and blood / total cfDNA samples were analysed using ddPCR assays, unless otherwise stated. All patients in which ddPCR was carried out had buffy coat and tumour also included as controls. Raw ddPCR results are not presented unless relevant and are summarised in appendix 7.6 (Table 49).

	Patient	Tumour Variant Detected	Number of Plasma Samples	Total cfDNA positive for tumour specific variant	
				YES/NO	No. Positive Samples
1	CRB62	YES	5	YES	4
2	CRB63	YES	5	NO	0
3	CRB64	YES	1	NO	0
4	CRB67	YES	2	YES	2
5	CRB69	YES	5	YES	3
6	CRB74	YES	6	NO	0
7	CRB75	YES	6	YES	2
8	CRB83	YES	1	YES	1
9	CRB84	YES	1	YES	1
10	CRB85	YES	1	YES	1
15	CRB86	NO	3		-
11	CRB89	YES	3	YES	3
16	CRB99	No tumour DNA	2		-
17	CRB108	DNA poor quality	2		-
14	CRB109	NO	2		-
12	CRB118	YES	1	YES	1
13	CRB121	YES	2	NO	0
	Total	YES = 13	48	YES= 9	18

Table 38 - **Summary of total cfDNA analysis for tumour specific variants for palliative patients**

5.6.1.1 Timelines showing utility of ctDNA for disease monitoring

All four timelines (CRB63, 69, 75 & 89, Figure 46-Figure 49) show dynamic changes in total cfDNA. In three of the patients initial chemotherapy treatment correlated with reduction in total cfDNA to undetectable levels (CRB62, 69 and 75). These responses were accompanied by prolonged progression free survival times of 24-68 weeks from the end of first chemotherapy (171 – 372 days). In two cases the drop to undetectable ctDNA was accompanied by a partial response. However, for CRB62 the CT showed stable disease following EOF chemotherapy despite a reduction in ctDNA which suggested a more significant response. The time to relapse of 24 weeks in this case, is more commonly seen in patients who have a partial response on their CT treatment (Webb et al., 1997, Ochenduszko et al., 2015). This therefore suggests that ctDNA levels are a good biomarker for tracking disease response to therapy and may give additional insight into disease activity than CT scans alone.

Patient CRB69 had disease progression at the time blood 3 was taken (Figure 47), which showed very high levels of ctDNA (2850 copies/m). Following EOX the CT scan showed a partial response to treatment and the ctDNA levels again declined to undetectable, both pointing to the fact that giving EOX was a good clinical decision. However, the patient then relapsed to the brain within 12 weeks of this partial response. Blood 5, taken shortly after this relapse was again positive at 121 copies/ml. It is somewhat surprising, given the finding in the first line treatment, that such a large reduction in ctDNA levels occurred but the patient relapsed shortly afterwards. However, the fact that ctDNA returns rapidly suggests that either the brain disease or the lymph node disease (or both) was producing ctDNA.

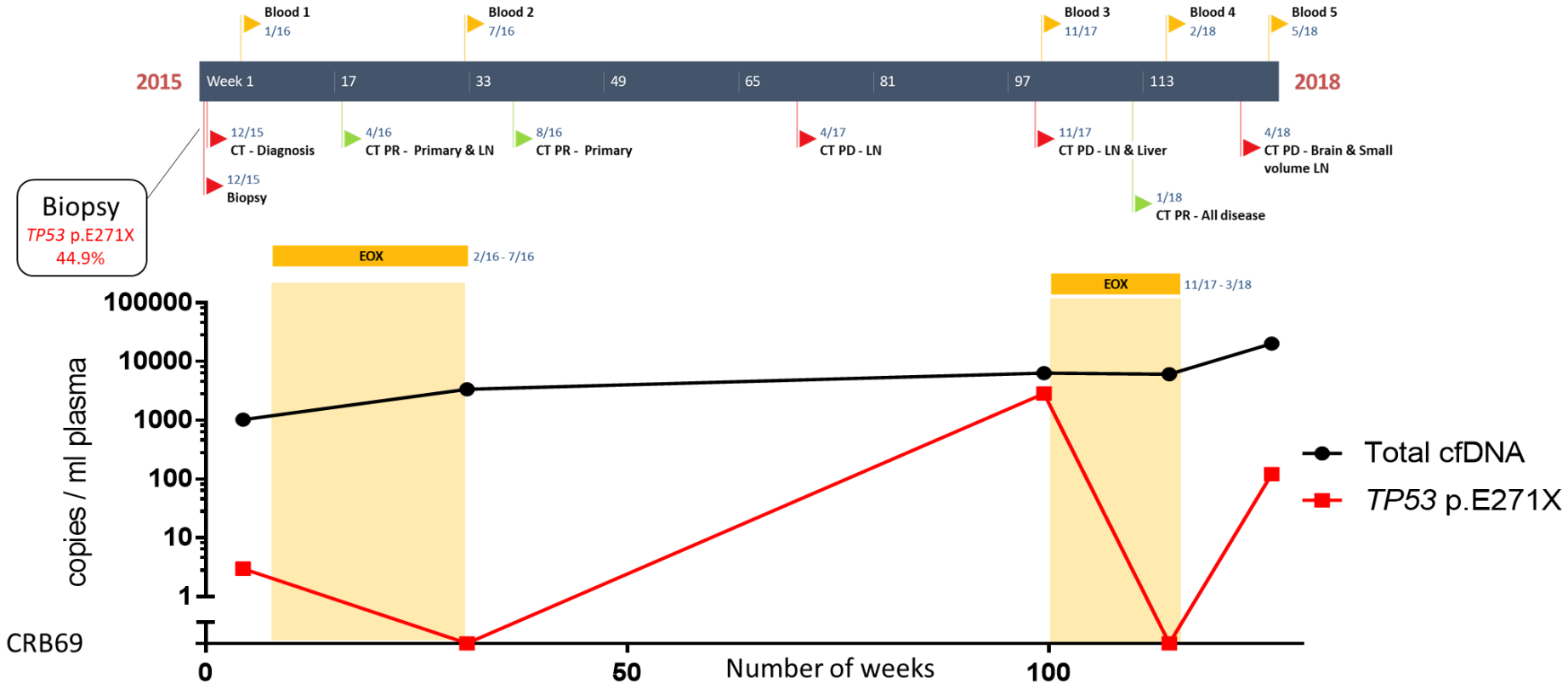
In all four patients ctDNA levels were highest just before death. This suggests that high ctDNA levels are a poor prognostic sign, which is not surprising given that rising levels are also associated with disease progression.

Patient CRB89 had stable disease and during the monitoring period and remained ctDNA positive throughout (Figure 49).

The timelines of CRB62 and 89 also suggest ctDNA levels can predict disease relapse / progression. For CRB62, blood sample 3 showed return of detectable ctDNA, preceding clinical confirmation of relapse by 1.5 weeks (10 days). Whereas for patient CRB89 a rise in ctDNA occurred 7.5 weeks (53 days) prior to disease progression being detected on the CT-scan. Unfortunately, patient CRB89 died 2 weeks after the CT scan before any further treatment could be initiated.

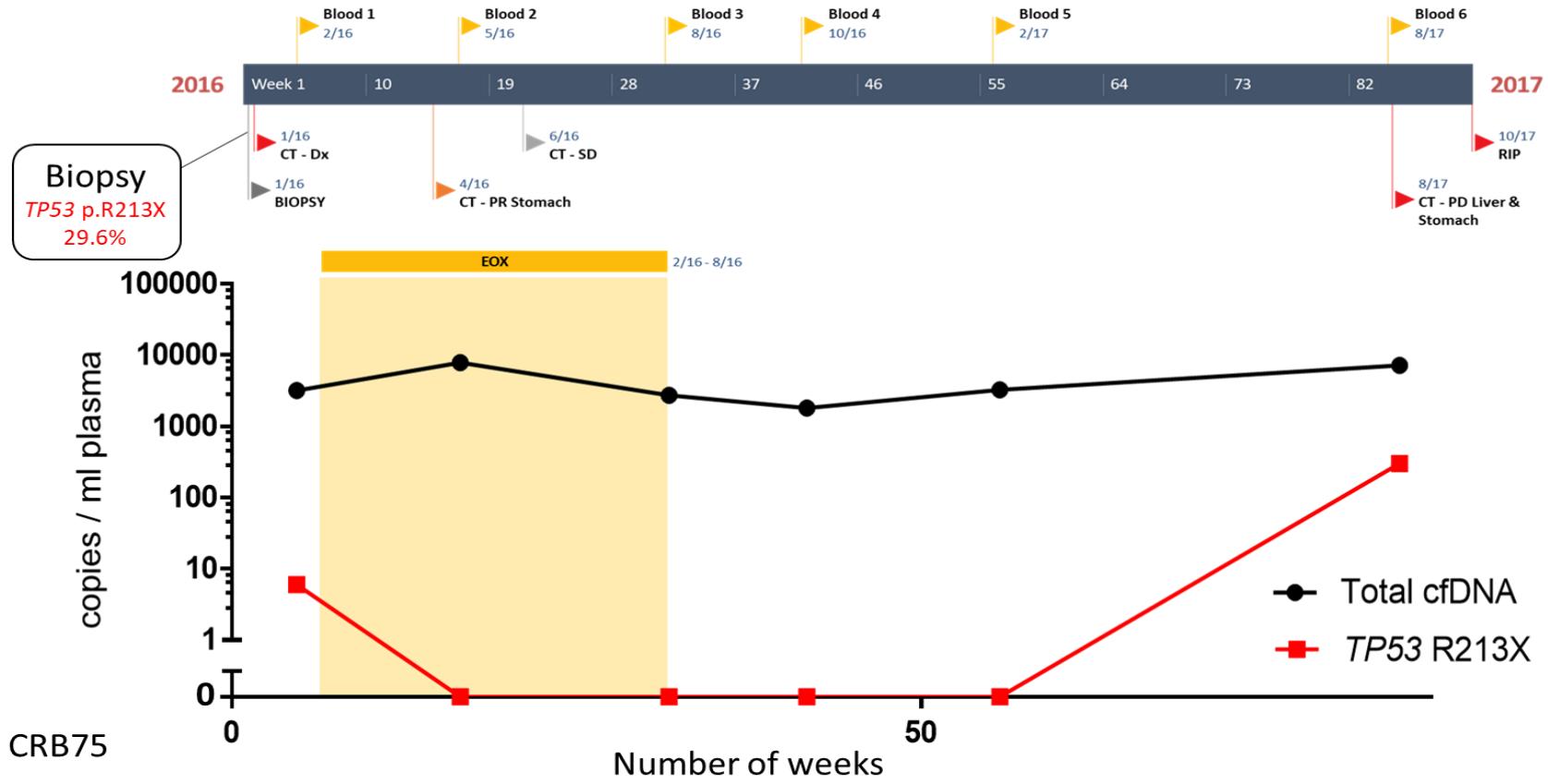
Blood Number		1	2	3	4	5
cfDNA copies /ml		1703	5581	2575	6866	7273
TP53 p.R280K	AF (%)	0.19	0	2.19	0.8	9.3
	Mutant copies/ml	32	0	56	54	676

Figure 46 - **Timeline for patient CRB62**. Showing clinical events in relation to total cfDNA copies/ml (black) and ctDNA as defined by detection of *TP53* p.R280K also shown as copies/ml (red) . The biopsy results are those from ddPCR experiments, although the peritoneal biopsy was also sequenced via the ampliseq GOA panel. SD = Stable disease, PD = progressive disease, PR = Partial response, CR = complete response, LN = Lymph node.



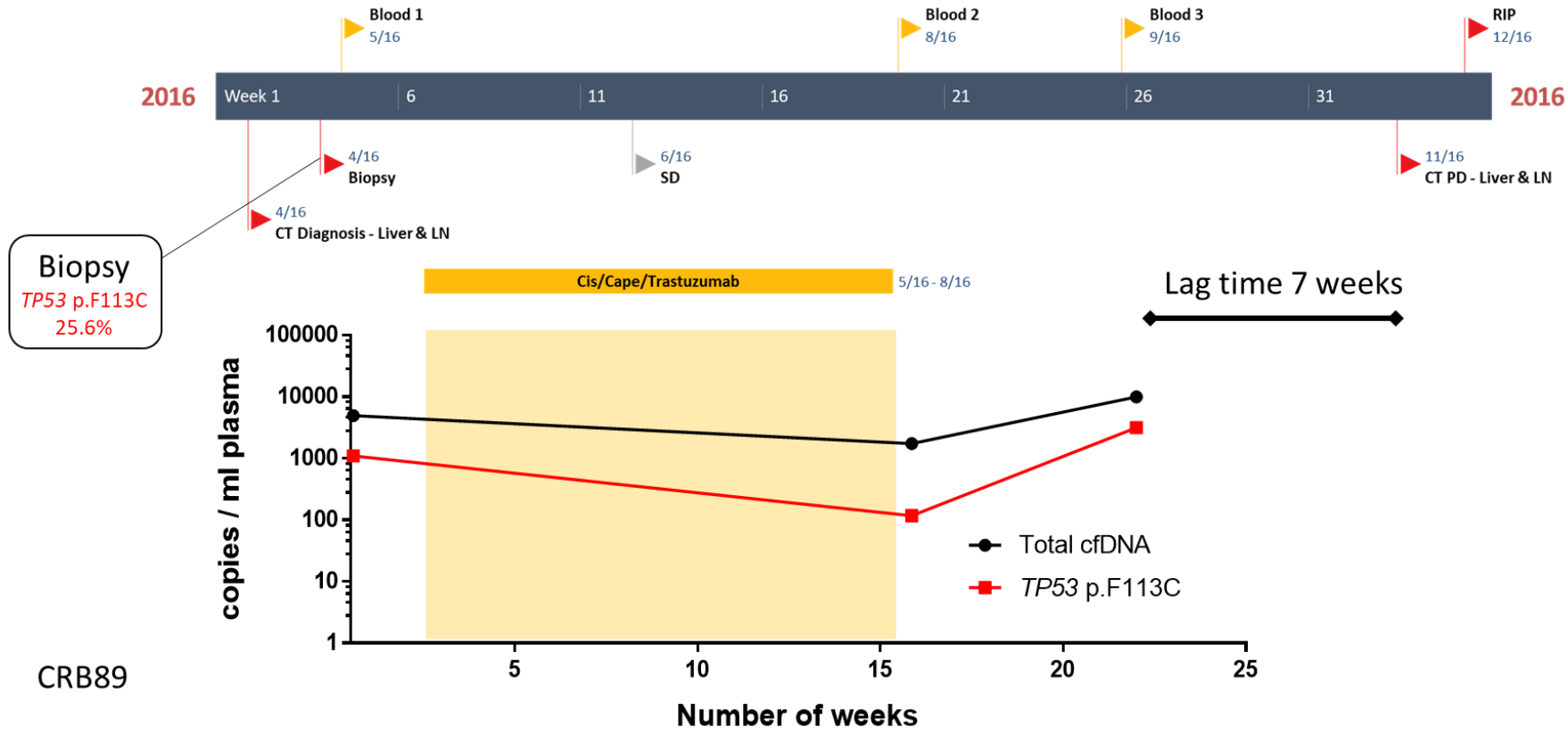
Blood Number		1	2	3	4	5
cfDNA copies /ml		1030	3393	6403	6100	20202
TP53 p.E271X	AF (%)	3.0	0	44.2	0	0.6
	Mutant copies/ml	3	0	2850	0	121

Figure 47 - **Timeline for patient CRB69**. Showing clinical events in relation to total cfDNA copies/ml (black) and ctDNA as defined by detection of *TP53* p.E271X also shown as copies/ml (red) . SD = Stable disease, PD = progressive disease, PR = Partial response, LN = Lymph node.



Blood		1	2	3	4	5	6
cfDNA copies /ml		3182	7879	2727	1818	3258	7172
TP53 p.R213X	AF (%)	0.19	0	0	0	0	4.2
	Mutant copies/ml	6	0	0	0	0	301

Figure 48 - **Timeline for patient CRB75.** Showing clinical events in relation to total cfDNA copies/ml (black) and ctDNA as defined by detection of *TP53* p.R213X also shown as copies/ml (red) . SD = Stable disease, PD = progressive disease, PR = Partial response.



Blood Number		1	2	3
cfDNA copies /ml		4924	1742	9924
TP53 p.F113C	AF (%)	22.3	6.7	31.6
	Mutant copies/ml	1098	117	3136

Figure 49 - **Timeline for patient CRB89.** Showing clinical events in relation to total cfDNA copies/ml (black) and ctDNA as defined by detection of *TP53* p.F113C also shown as copies/ml (red) . Lag time between rising ctDNA levels and relapse shown as solid bar. SD = Stable disease, PD = progressive disease, LN = Lymph nodes

5.6.1.2 ctDNA negative metastatic patients

There were 4 patients with a detectable mutation in tumours DNA > 5% VAF, that was not detectable in corresponding total cfDNA.

Patient CRB63, had a *TP53* p.R282W (38.1% VAF) mutation in tumour DNA of a suspected lung metastasis. However, neither sequencing nor ddPCR of the plasma DNA samples detected this mutation, meaning ctDNA was undetectable. This patient had a different medical history to the other metastatic patients in this study, in that their only site of metastatic disease, a lung tumour, had been resected (Figure 50) prior to blood sampling. Following metastectomy they went on to receive radical chemo-radiotherapy, with the hope of eradicating any localised disease. The patient understood that in the setting of metastatic disease this would not be a curative treatment, but may prolong survival. The *TP53* p.R282W mutation was absent by NGS from the primary tumour, which was unexpected. The mutation was confirmed to be present in the lung tumour by ddPCR with a VAF of 35.1% and in DNA from the primary biopsy the *TP53* mutation was detected at a low VAF of 0.18% (Figure 50). Thus it was not clear if this was a case of two independent primaries or of metastatic disease with an emergent *TP53* mutation.

To further investigate this, the tumour samples were analysed as part of a separate study called 'primings'. This study investigates patients with known previous cancers who have developed a further tumour, where it is uncertain if this new tumour is a new primary cancer or a metastatic deposit of the previous cancer. The project takes DNA from the known primary tumour tissue and the new tumour tissue of uncertain origin and sequences a large number of known cancer genes using the ampliseq cancer hotspot version 2 panel. It then compares the number of shared genetic aberrations between the two tumours to predict whether it is a case of a new primary or metastatic deposit to inform clinical decision making. The result of this was is shown in Table 39. There was an undetected mutation in the gene *CDKN2A* p.R80X, that was shared between the two tumours indicating that this

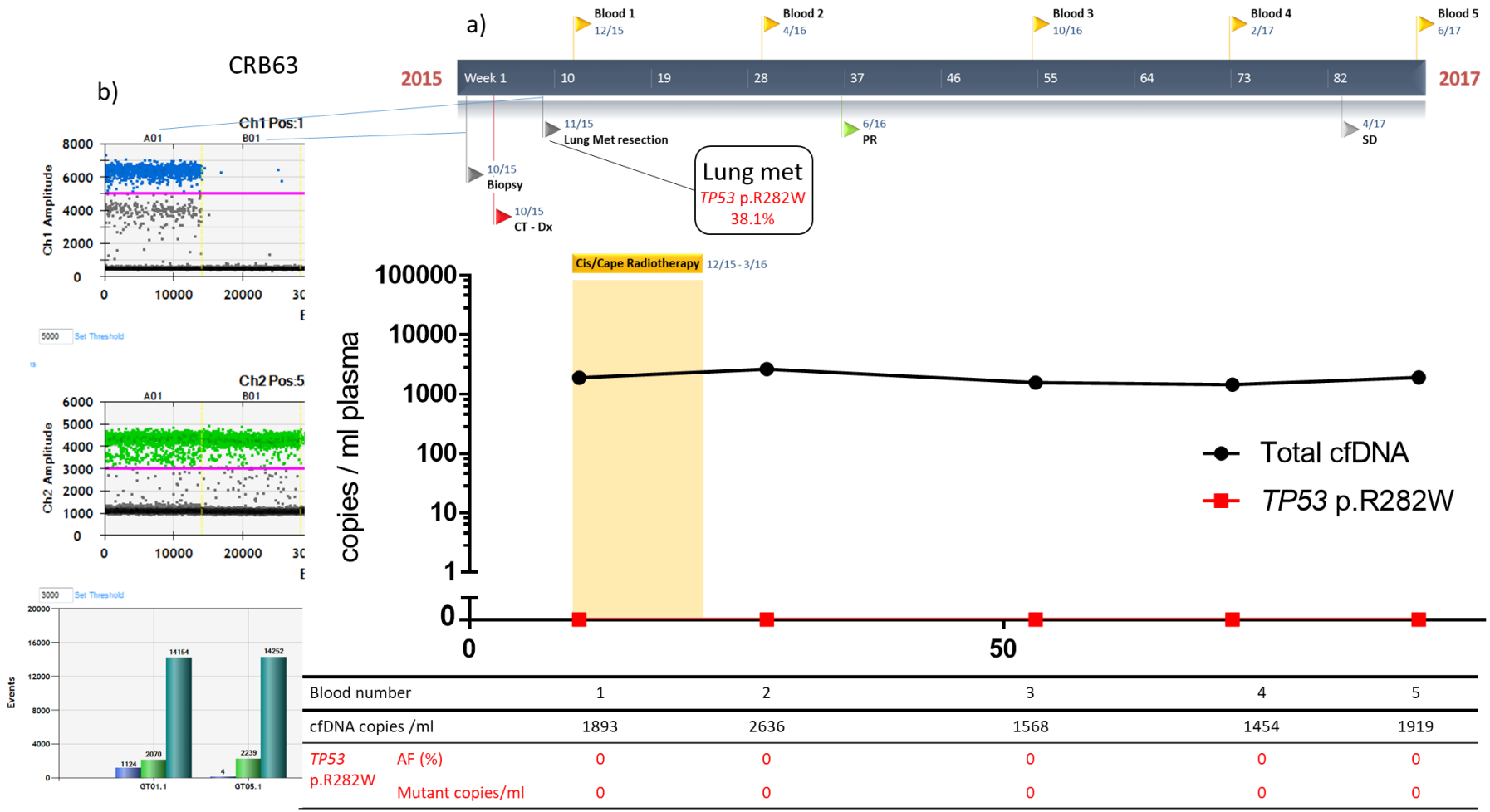
was indeed a case of resected metastatic disease, with an emergent *TP53* mutation which was detected at low levels in the primary tumour DNA by ddPCR.

Ion AmpliSeq™ Cancer Hotspot v2									
Position	**Ref**	**Alt**	**Gene**	**COSMIC ID**	**Primary Tumour (Mutation Frequency)**	**Second Tumour (Mutation Frequency)**	**CDS Mutation**	**AA Mutation**	**Diagnosis**
chr17:7577094	G	A	***TP53***	COSM99925	**0%**	**35.6%**	c.844C>T	p.R282W	**Metastasis**
chr9:21971120	G	A	***CDKN2A***	COSM12475	**36.7%**	**30.9%**	c.238C>T	p.R80X	

Table 39 – **Results of primings analysis of CRB63 using the ampliseq cancer hotspot version 2 panel.** Used with permission of Luke Martinson PhD student.

Unfortunately it was not possible to design a ddPCR assay for the *CDK2NA* p.R80X mutation as the SNV lies in an extremely GC rich region of the gene, and therefore it was not possible to check if ctDNA carrying this mutation was present.

These finding suggest that the lack of detection of ctDNA is because the metastatic disease was well controlled. The patient has survived for a long time without a recurrence, (134 weeks, 943 days), and it is possible, albeit unlikely, that the patient is cancer free. As the *TP53* p.R282W is at high levels in the metastatic site it would be expected to be present if any ctDNA was present. However, undertaking sequencing of total cfDNA for the *CDK2NA* p.R80X mutation could yet reveal presence of ctDNA.



Blood number		1	2	3	4	5
cfDNA copies /ml		1893	2636	1568	1454	1919
TP53 p.R282W	AF (%)	0	0	0	0	0
	Mutant copies/ml	0	0	0	0	0

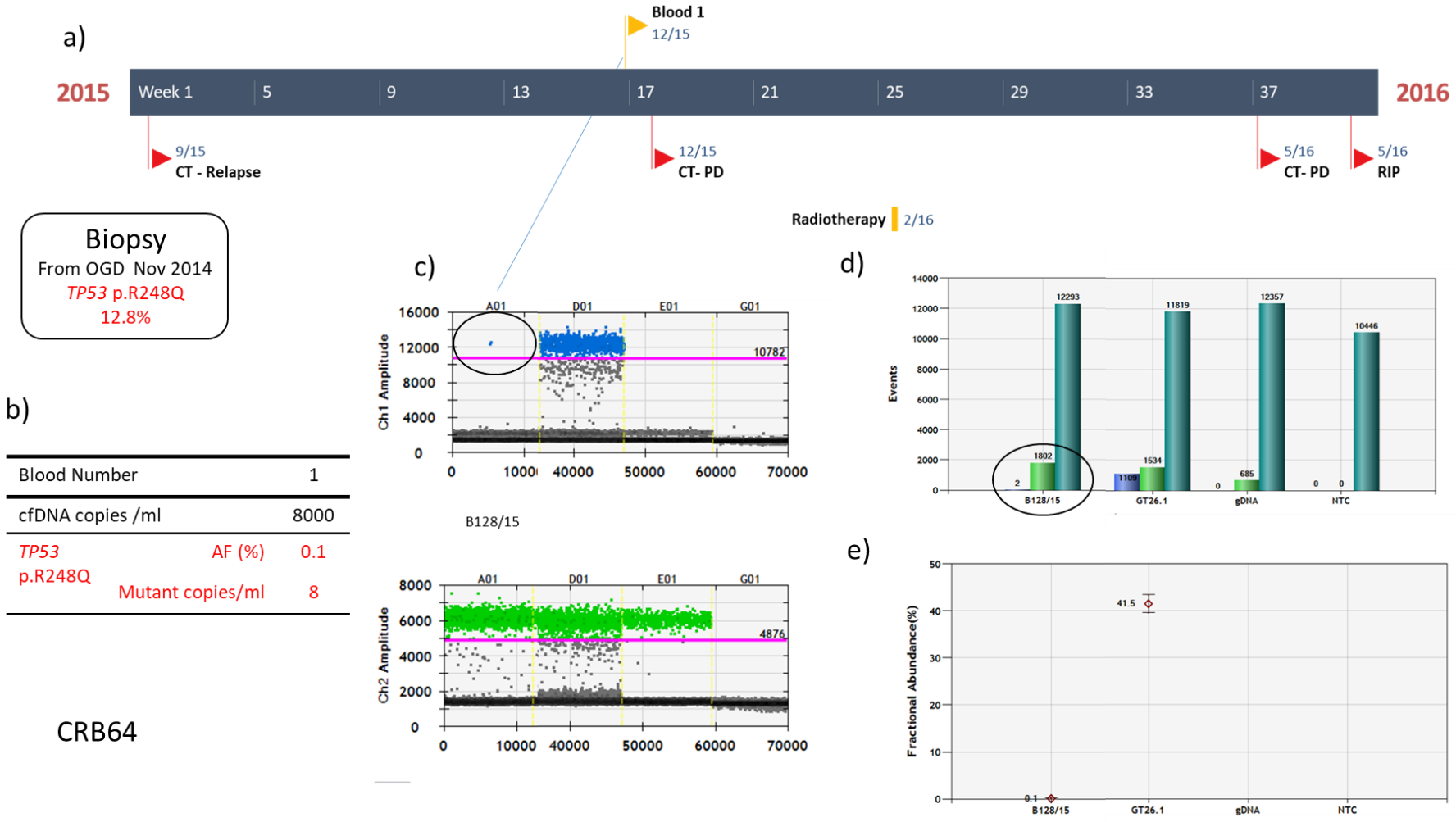
Figure 50 - **Timeline for patient CRB63.** a) Showing clinical events in relation to total cfDNA copies/ml (black) and ctDNA as defined by detection of *TP53* p.R282W also shown as copies/ml (red). b) ddPCR results for *TP53* p.R282W of the original biopsy (second column GT05.1) and the resected lung metastasis (first column, GT01.1). Blue dots represent droplets positive for the mutation, and green droplets those containing only wildtype (WT) DNA. SD = Stable disease, PD = progressive disease, PR = Partial response, CR = complete response, LN = Lymph node.

Patient CRB64 had a single blood test, B128/15 (Figure 51) and was recruited after referral to medical oncology, but was deemed unfit for surgery. Tumour DNA sequencing detected a *TP53* p.R248Q mutation (12.3% VAF). Total cfDNA ddPCR for the *TP53* p. R248Q mutation was equivocal. Only two positive droplets were seen in the plasma despite good loading of the reaction with >1800 wildtype droplets (Figure 51) and was repeated with the same result. Within our lab a threshold of 3 positive droplets must be seen to call the sample positive. However, with 2 positive droplets in replicate it is more likely this is a real reading, with very low levels of ctDNA. The patient survived for 22 weeks without systemic chemotherapy, but did receive palliative radiotherapy to the oesophagus.

Patient CRB74, did not undergo surgery with curative intent, despite initially being selected for potentially curative treatment and neoadjuvant chemotherapy. Following neoadjuvant chemotherapy lymph node disease was found to have shrunk outside of the field that could be surgically removed, and therefore curative surgery could not be undertaken. Tumour DNA sequencing detected a *TP53* p.S241A mutation (34.1%). However, despite having widespread lymph node disease at diagnosis and having progressive disease later in his treatment, no evidence of this same mutation was found in the blood, and all 6 of the patient's blood tests were ctDNA negative (Figure 52). Total cfDNA levels fluctuated at 1400-4000 copies/ml.

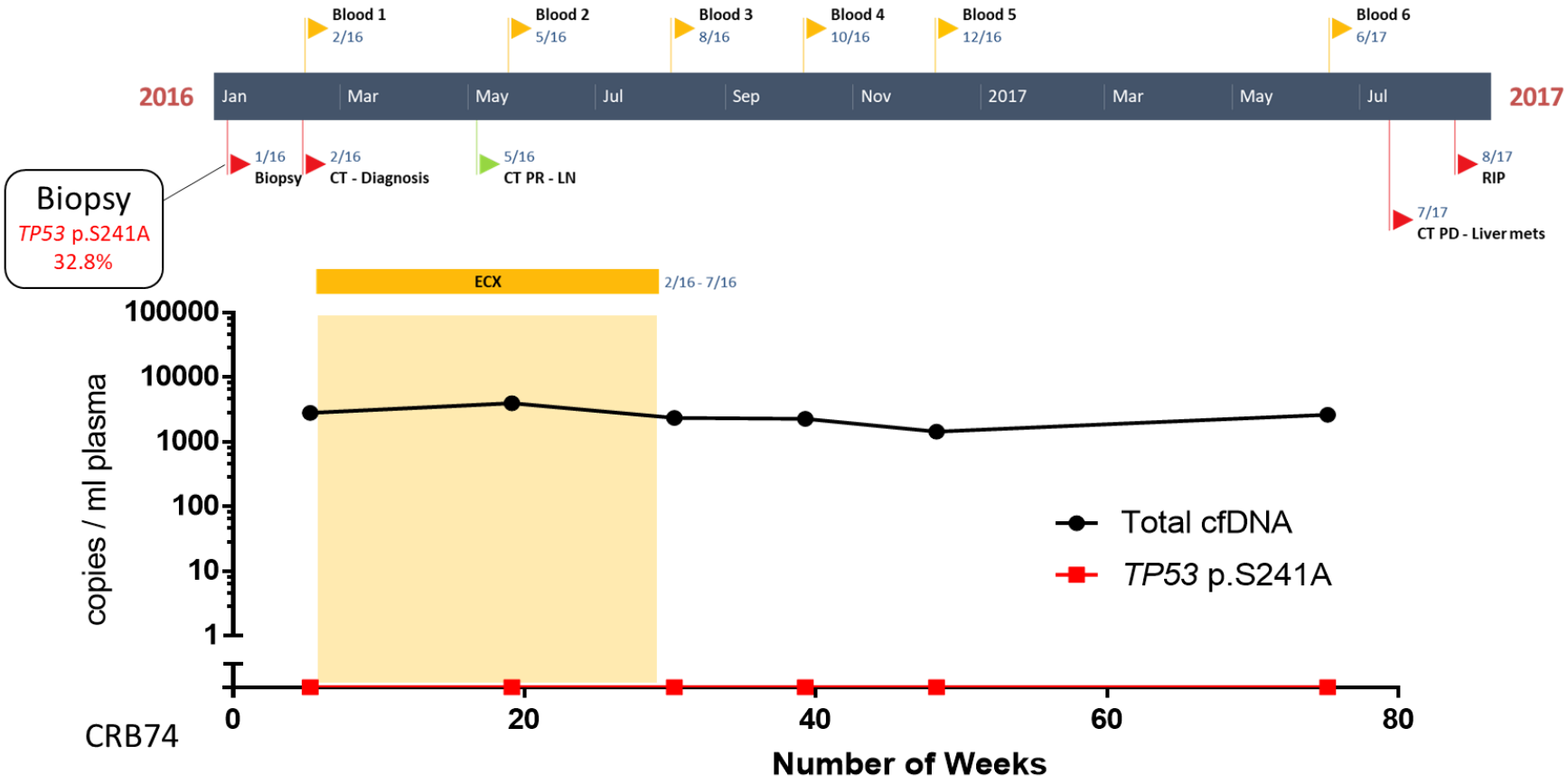
Patient CRB121 had a *TP53* p.V122Dfs26X mutation (44.9%). No blood was available pre-treatment and the patient had localised progressive disease during her follow-up. Both blood samples were negative (Figure 53).

In summary, these four patients had no evidence of ctDNA in the plasma based upon tracking of a single mutation detected in the primary tumour. It would be interesting to analyse the tumour and plasma of these patients by WES to look for other drivers of the cancer.



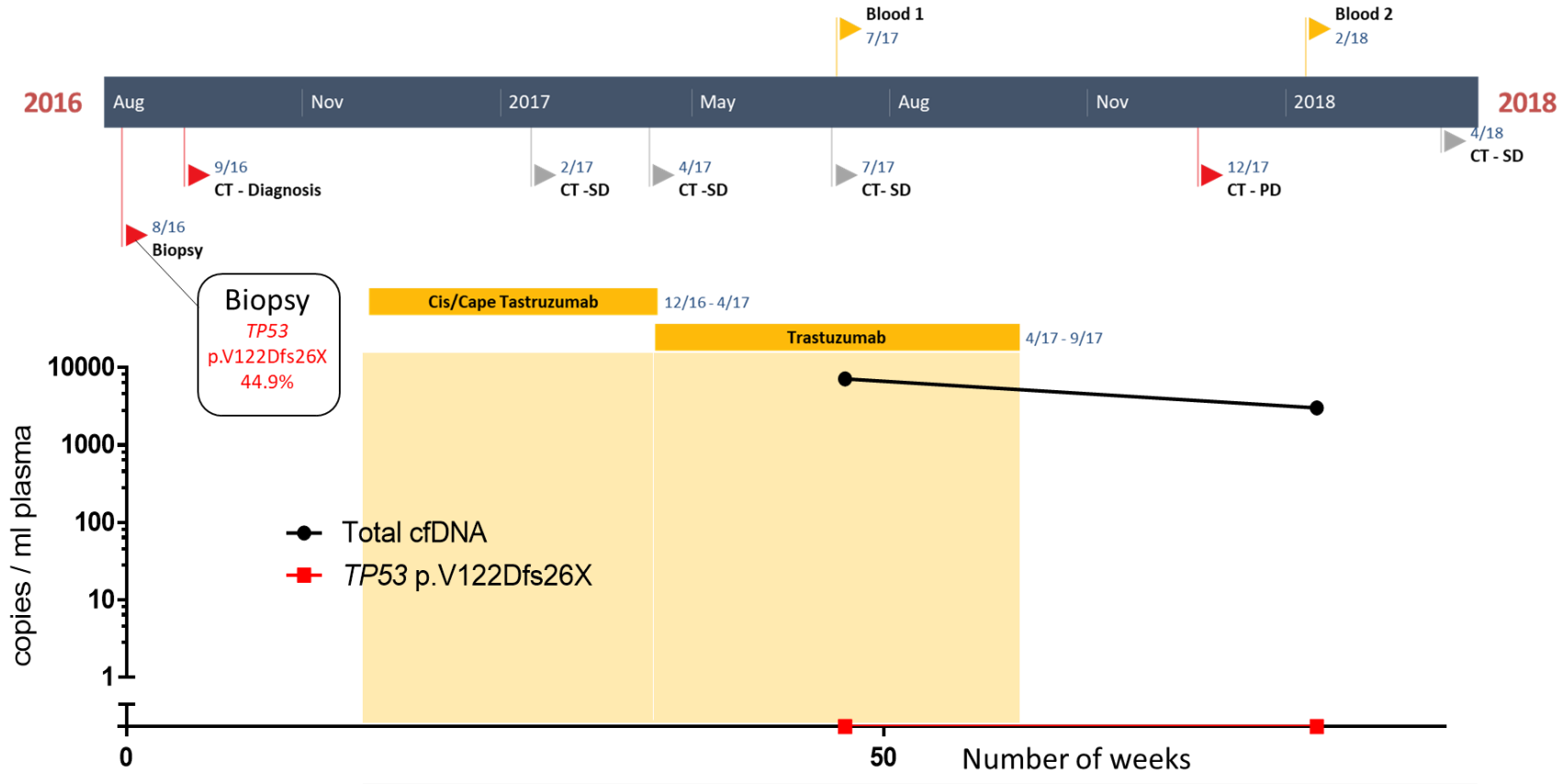
Blood Number		1
cfDNA copies /ml		8000
TP53 p.R248Q	AF (%)	0.1
	Mutant copies/ml	8

Figure 51 - **Timeline for patient CRB64.** a) Showing clinical events, PD = progressive disease. b) Summary table of total cfDNA copies/ml c) ddPCR results for *TP53* p.R248Q for the only blood sample in lane A01, B128/15 showing just two positive droplets. d) The total number of events, and positive droplets for mutant and wildtype DNA. Oval highlighting results for B128/15. e) fractional abundance GT26.1 = positive control DNA, gDNA = genomic DNA, NTC = non-template control.



Blood Number		1	2	3	4	5	6
cfDNA copies /ml		2810	3939	2340	2265	1439	2624
TP53 p. S241A	AF (%)	0	0	0	0	0	0
	Mutant copies/ml	0	0	0	0	0	0

Figure 52 - **Timeline for patient CRB74.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.S241A mutation not detected in the total cfDNA by ddPCR (red). PR = Partial response, PD = Progressive disease, LN = Lymph node.



Blood Number		1	2
cfDNA copies /ml		7175	3030
TP53 p.V122Dfs26X	AF (%)	0	0
	Mutant copies/ml	0	0

Figure 53 - **Timeline for patient CRB121.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.V122Dfs26X mutation not detected in the total cfDNA by ddPCR (red). PR = Partial response, PD = Progressive disease, LN = Lymph node.

5.6.2 GOA patients who received surgery with curative intent.

23 of 27 patients initially selected for neoadjuvant chemotherapy went on to have curative surgery, these 23 patients are hereon described as patients treated with 'curative intent'. One patient did not have tumour DNA of suitable quality for sequencing. Following sequencing with the GOA ampliseq panel, two patients (CRB78, CRB94) had no detectable mutation and one (CRB102) had a low frequency SNV (<5% VAF) and therefore these patients are not described further in this section. In total 19 of the 23 patients had a detected SNV >5% VAF suitable for taking forward for total cfDNA analysis by NGS and ddPCR and 7 of the 19 patients (36.8%) had the same mutation detected in total cfDNA, indicating the presence of ctDNA (Table 40).

	Patient	Tumour Variant Detected	Number of Plasma Samples	Total cfDNA positive for tumour specific variant	
				YES/NO	No. Positive Samples
1	CRB55	YES	4	YES	3
2	CRB58	YES	8	YES	4
3	CRB59	YES	5	NO	0
4	CRB61	YES	3	YES	2
5	CRB70	YES	3	NO	0
6	CRB71	YES	4	NO	0
7	CRB78	NO	3		-
8	CRB80	YES	4	NO	0
9	CRB82	YES	3	NO	0
10	CRB88	YES	4	NO	0
11	CRB94	NO	3		-
12	CRB95	YES	3	NO	0
13	CRB102	NO	5		-
14	CRB105	YES	3	YES	1
15	CRB107	YES	2	YES	2
16	CRB110	YES	3	NO	0
17	CRB114	DNA poor quality	1		-
18	CRB115	YES	3	NO	0
19	CRB116	YES	5	YES	1
20	CRB117	YES	3	NO	0
21	CRB137	YES	2	YES	1
22	CRB144	YES	1	NO	0
23	CRB145	YES	1	NO	0
	Total	YES = 19	76	YES = 7	14

Table 40 - **Summary of total cfDNA analysis for tumour specific variants of all curative intent patients**

Total cfDNA samples were analysed via ddPCR for the highest frequency mutation in their respective tumour DNA samples unless otherwise stated. Buffy coat (lymphocyte) DNA for each patient was analysed by ddPCR to confirm the mutation was not germline in origin. Raw ddPCR results are not presented unless relevant.

8 of the 19 patients who were amenable to ctDNA profiling relapsed in the course of the study. 4 of the 8 had ctDNA detected at relapse and a further 2 had ctDNA detected prior to relapse. 2 relapse patients had no detectable ctDNA at any time point during their follow-up (Table 41). The results and timelines of these 8 patients are described in further detail.

	Patient	Relapse?	Tumour DNA mutation			Number of Plasma Samples	Total cfDNA positive for tumour specific variant	
		YES/NO	Gene	Mutation	VAF (%)		YES/NO	No. Positive Samples
1	CRB55	YES	*TP53*	R196X	71.9	4	YES	3
			PIK3CA	I338N	37.2			
2	CRB58	YES	*KRAS*	G12D	37.4	8	YES	4
3	CRB59	NO	*TP53*	C277F	62.2	5	NO	0
4	CRB61	YES	*TP53*	R175H	29.5	3	YES	2
5	CRB70	NO	*TP53*	R213X	5.2	3	NO	0
6	CRB71	YES	*TP53*	R248Q	35.3	4	NO	0
7	CRB80	YES	*TP53*	R175H	6.4	4	NO	0
8	CRB82	NO	*TP53*	R306X	19.6	3	NO	0
9	CRB88	NO	*TP53*	E286K	36.4	4	NO	0
10	CRB95	NO	*TP53*	P151R	53.2	3	NO	0
			ARID1A	Y1319X	44.2			
11	CRB105	YES	*TP53*	C135F	56.0	3	YES	1
12	CRB107	YES	*TP53*	Q192X	61.5	2	YES	2
13	CRB110	NO	*TP53*	C176F	27.8	3	NO	0
14	CRB115	NO	*TP53*	H179R	6.0	3	NO	0
			RHOA	A44V	9.0			
15	CRB116	YES	*TP53*	G266V	58.2	5	YES	1
16	CRB117	NO	*TP53*	R213X	20.9	3	NO	0
17	CRB137	NO	*TP53*	T125T	88.7	2	YES	1
18	CRB144	NO	*TP53*	S215R	57.5	1	NO	0
19	CRB145	NO	*TP53*	E286K	34.0	1	NO	0
			PIK3CA	E545K	6.4			
		YES=8				Total = 64	YES= 7	Total = 14

Table 41 **– Summary of relapse status of all curative intent patients with a trackable tumour DNA variant.**

5.6.2.1 Cancer tumour evolution

Patient CRB55 gave significant insight into tumour evolution (Figure 54). The primary tumour had 2 mutations, *TP53* p.R196X (71.9 % VAF), confirmed by ddPCR and a subclonal *PIK3CA* p.I338N (37.2% VAF) confirmed by sequencing with the Acquires panel (section 5.5.1.). Four serial total cfDNA samples were analysed by both ddPCR of the *TP53* p.R196X variant and GOA panel sequencing due to the absence of a ddPCR assay for *PIK3CA* p.I338N.

Sequencing of total cfDNA detected the *TP53* p.R196X mutation alone, the *PIK3CA* p.I338N mutation was absent. This suggests that the metastatic disease in this patient is derived from a clone of the original tumour containing cells carrying the *TP53* variant but not the *PIK3CA* one.

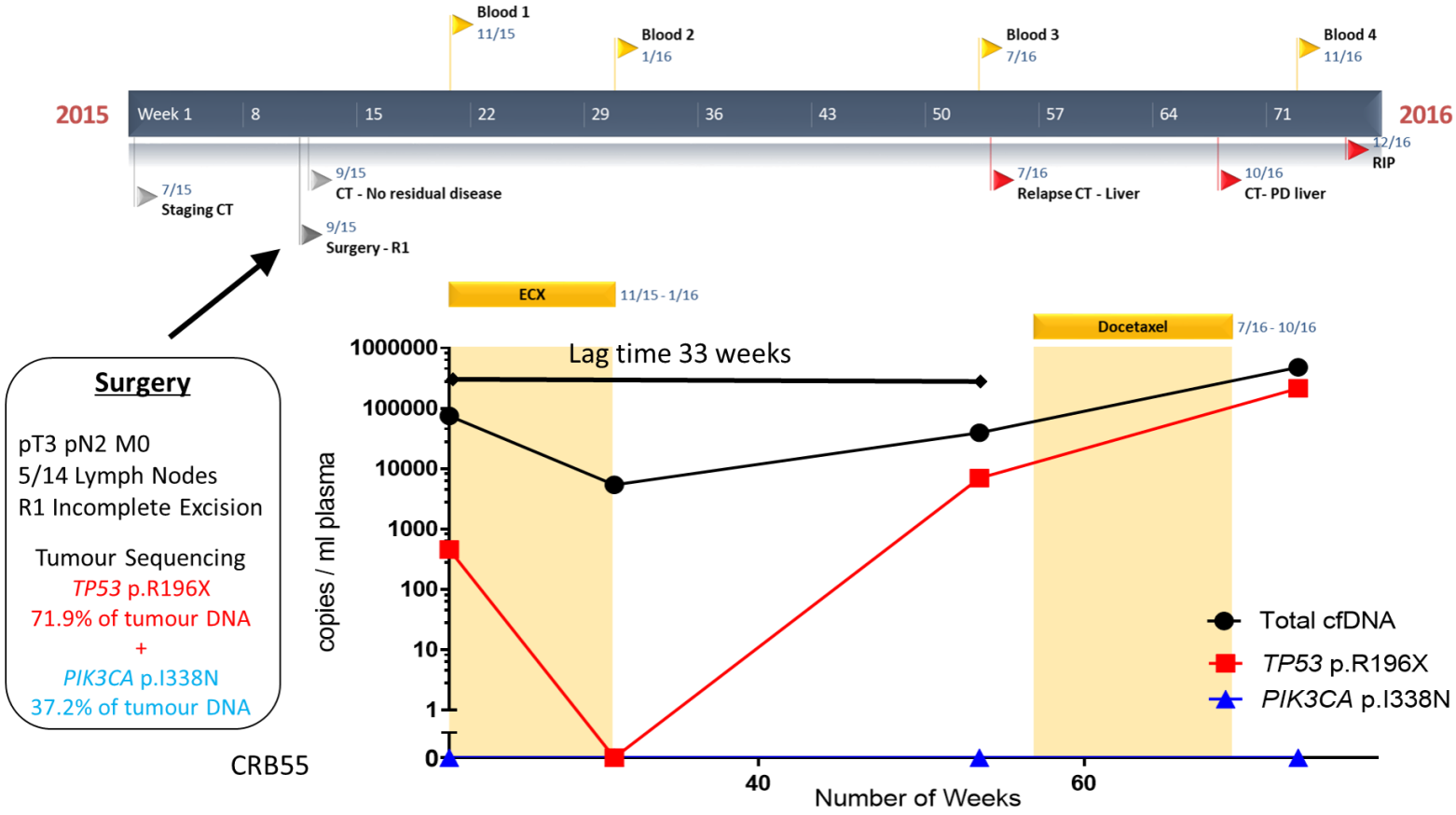
5.6.2.2 Positive ctDNA and detection of minimal residual disease

Here I describe the timelines of the 6 patients who had positive ctDNA and went onto relapse.

The first, post-operative blood sample for CRB55 (blood 1), taken 8 weeks (56 days) post surgery, detected the *TP53* p.R196X variant at a low level, when the patient was well and apparently disease free (Figure 54). The ctDNA became undetectable with adjuvant chemotherapy (blood 2). The presence of ctDNA post surgery is evidence for the presence of minimal residual disease. In keeping with the prediction of presence of minimal residual disease the patient's disease relapses 43 weeks (298 days) following surgery when ctDNA is again detectable (blood3). ctDNA then continued to rise following relapse (blood 4) despite docetaxel chemotherapy, suggesting non-response of disease to therapy. In keeping with this finding progressive disease was seen on CT scanning, therefore confirming the ability of ctDNA to predict non response to treatment in this patient. The final blood test just before death has one of the highest ctDNA levels in the project (215,360 copies/ml) in keeping with ctDNA being a poor prognostic sign.

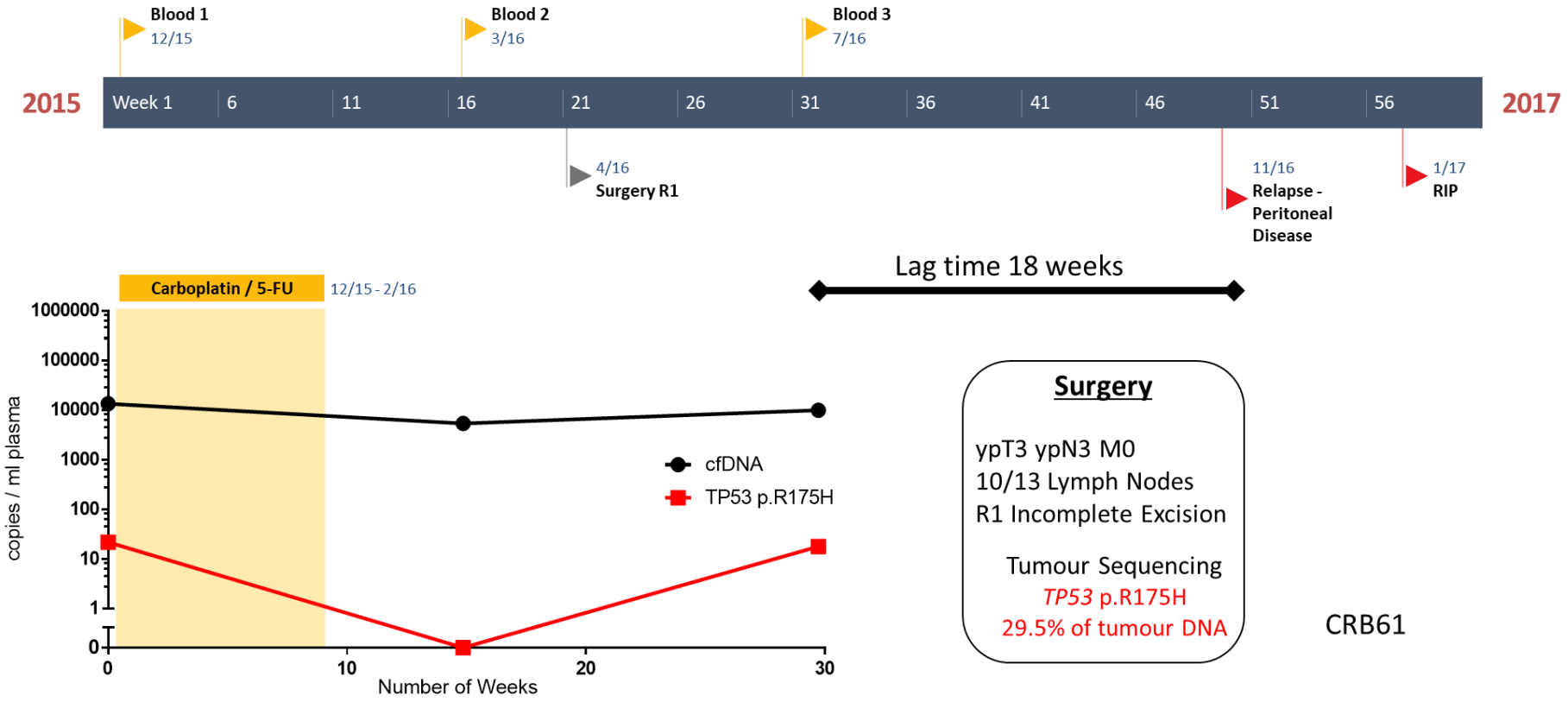
Patient CRB61 had two serial blood samples whilst on neoadjuvant chemotherapy. They also relapsed following curative surgery (Figure 55). A *TP53* p.R175H mutation (29.5% VAF) was detected in the tumour DNA. The initial blood showed low levels of ctDNA, which became undetectable following neoadjuvant chemotherapy and was again detectable following surgery in blood 3, with a VAF of 0.18%, 18 copies/ml. The patient did not receive adjuvant chemotherapy and relapsed 29 weeks (203 days) after surgery. They died from complications of disease before a relapse blood sample could be taken.

These two cases present evidence that post-operative ctDNA heralds disease relapse, and may indicate the presence of minimal residual disease. Both patients relapsed a number of weeks following their positive post-operative sample at 33 and 18 weeks (233 and 128 days) respectively. This suggests a significant window of opportunity for therapeutic intervention.



Blood Number		1	2	3	4
cfDNA copies /ml		75303	5454	39469	481848
TP53 p.R196X	AF (%)	0.6	0	17.8	43.1
	Mutant copies/ml	459	0	7025	215360
PIK3CA p.1338N	AF (%)	0	0	0	0
	Mutant copies/ml	0	0	0	0

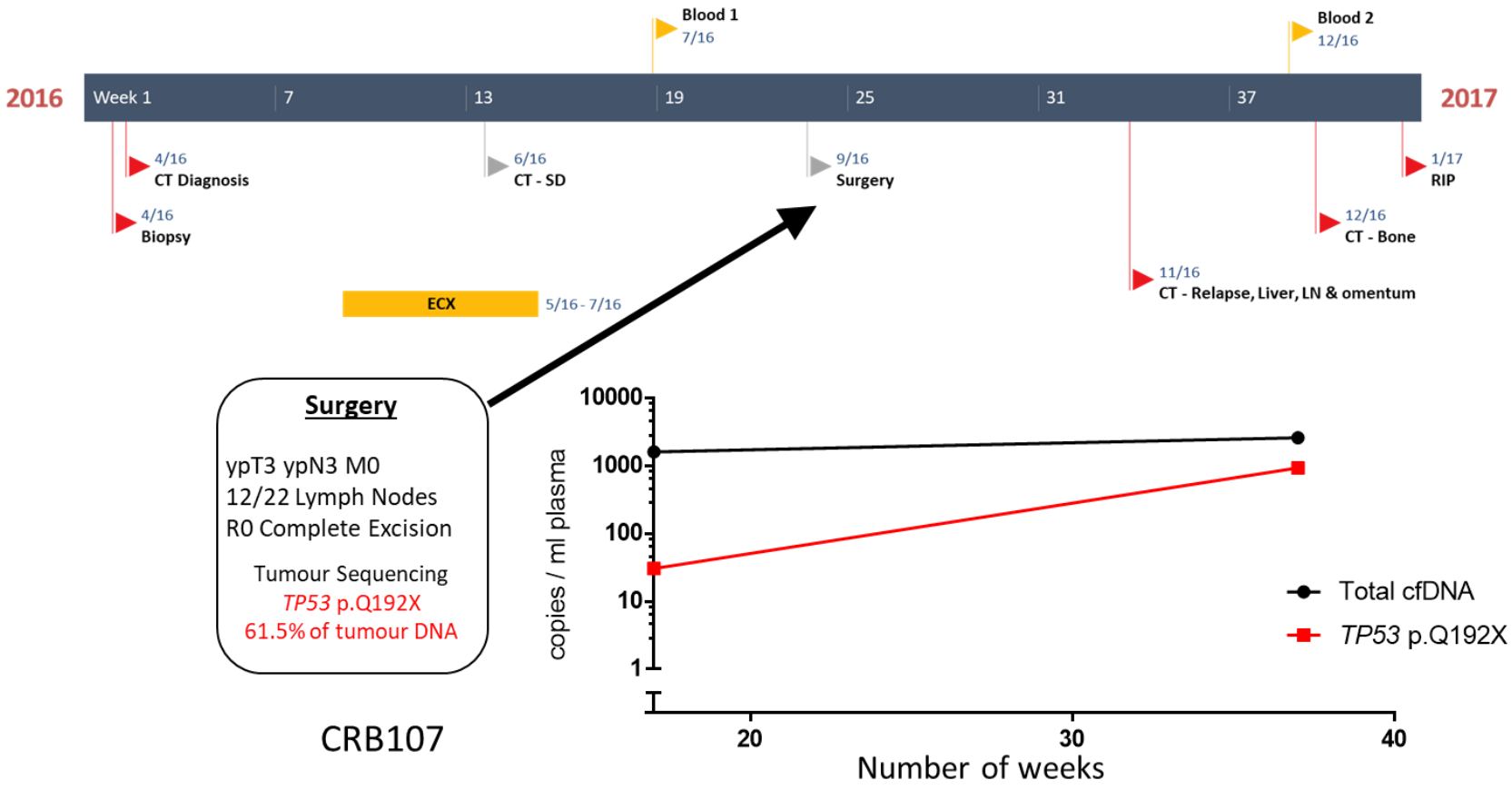
Figure 54 - **Timeline for patient CRB55.** Showing clinical events in relation to totalcfDNA copies/ml (black) and ctDNA as defined by detection of *TP53* p.R196X (by ddPCR) shown as copies/ml (red). Note *PIK3CA* p.I1338N was not detected in the total cfDNA. Lag time between first detectable ctDNA positive blood and relapse shown as solid bar. PD = progressive disease.



1	2	3		Blood number
13469	5378	9999		cfDNA copies /ml
0.16	0	0.18	AF (%)	*TP53* p.R175H
22	0	18	Mutant copies/ml	

Figure 55 - **Timeline for patient CRB61.** Showing clinical events in relation to total total cfDNA copies/ml (black) and ctDNA as defined by detection of TP53 p.R175H (by ddPCR) shown as copies/ml (red). Lag time between first detectable ctDNA positive blood and relapse shown as solid bar.

Patient CRB107 had 2 serial blood samples taken. The first was taken post neoadjuvant chemotherapy and prior to surgery, the second 5 months later after disease relapse. Relapse occurred 10 weeks (71 days) after surgery (Figure 56). The tumour was positive for a *TP53* p.Q192X mutation (61.5% VAF) and ctDNA was detected in both blood samples. This was the only patient where the blood sample was positive for ctDNA going into surgery, and it is likely that if a post-operative blood sample had been available it would have been positive, given the high level of ctDNA in the second blood sample (927 copies/ml). Given the findings of the pre surgery blood for CRB107 and that ctDNA appears to be a marker for active disease, it is not surprising that this patient relapsed 10 weeks after surgery, and died before a follow-up blood sample could be taken.

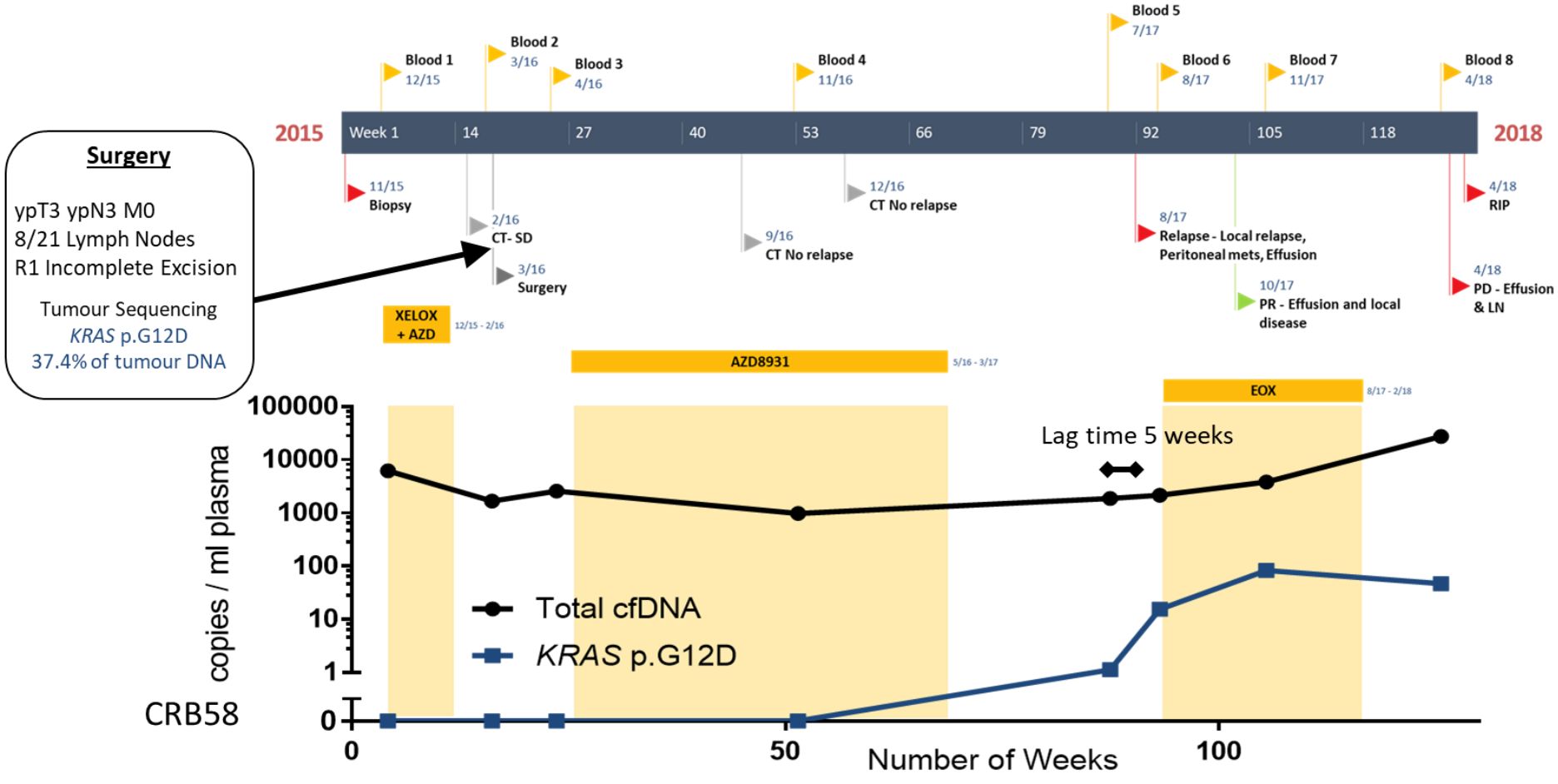


Blood number		1	2
cfDNA copies /ml		1590	2527
TP53 p.Q192X	AF (%)	1.9	36.0
	Mutant copies/ml	30	927

Figure 56 - **Timeline for patient CRB107**. Showing clinical events in relation to total cfDNA copies/ml (black) and ctDNA as defined by detection of TP53 p.Q192X (by ddPCR) shown as copies/ml (red). No pre-relapse blood was available. SD = stable disease

Patient CRB58 was monitored over more than 2 years through 8 serial blood samples, 2 of which were taken prior to surgery. A *KRAS* p.G12D mutation (37.4% VAF) was detected in the primary tumour and confirmed in the initial OGD biopsy. Despite high levels of total cfDNA in blood sample 1 (Figure 57), no mutation was detected in the blood at diagnosis or during follow-up until 3 weeks (21 days) before relapse. The patient underwent treatment in a clinical trial with drug AZD8931 now called Sapitibin, which is an epidermal growth factor receptor (EGFR) family (pan-erbB) tyrosine kinase inhibitor (Mu et al., 2014). They received AZD8931 for a full year. Following completion of the therapy the patient noticed a florid skin rash occurring. A similar skin rash had occurred before diagnosis and resolved during therapy with curative intent, suggesting that this may be a paraneoplastic event (Abreu Velez and Howard, 2010, Ge et al., 2014). Blood sample 5, taken at the time of recurrence of the skin rash showed two mutant droplets on ddPCR (circled on Figure 58) with a VAF of 0.06%, 1 copy/ml plasma. This result was confirmed by repeat of the experiment with more input DNA. The patient had a confirmed relapse of local and metastatic disease 3 weeks later and the next blood sample, taken 1 week after relapse (4 weeks from blood 5) definitively picked up the *KRAS* p.G12D mutation with a VAF of 0.72%, 15 copies/ml. This proximity of definitive relapse to the equivocal finding in blood 5, suggests that the result in sample 5 was indeed a low positive.

Patient CRB58 had a partial response on CT following relapse as shown by reduction in the pleural effusion and local disease on CT, but the ctDNA level continued to rise during the same period (Blood 7, Figure 57). The patient died from liver failure caused by disease obstructing the biliary tree, 10 weeks (63 days) following completion of therapy, in keeping with a minimal or short lived response to treatment. The final blood taken just before death (Blood 8) shows a significant decline in the VAF of the *KRAS* mutation but a significant increase in total cfDNA overall, potentially caused by high levels of total cfDNA from normal tissue, such as inflammation due to liver failure, and therefore overall there is only a slight decrease in *KRAS* levels as measured by copies/ml.



Blood number		1	2	3	4	5	6	7	8
cfDNA copies /ml		6212	1666	2596	984	1874	2163	3809	27474
KRAS p.G12D	AF (%)	0	0	0	0	0.06	0.72	2.5	0.17
	Mutant copies/ml	0	0	0	0	1	15	95	48

Figure 57 - **Timeline for patient CRB58**. Showing clinical events in relation to total cfDNA copies/ml (black) and ctDNA as defined by detection of *KRAS* p.G12D (by ddPCR) shown as copies/ml (blue). Lag time between first detectable ctDNA positive blood and relapse shown as solid bar. SD = Stable disease, PD = progressive disease, PR = partial response.

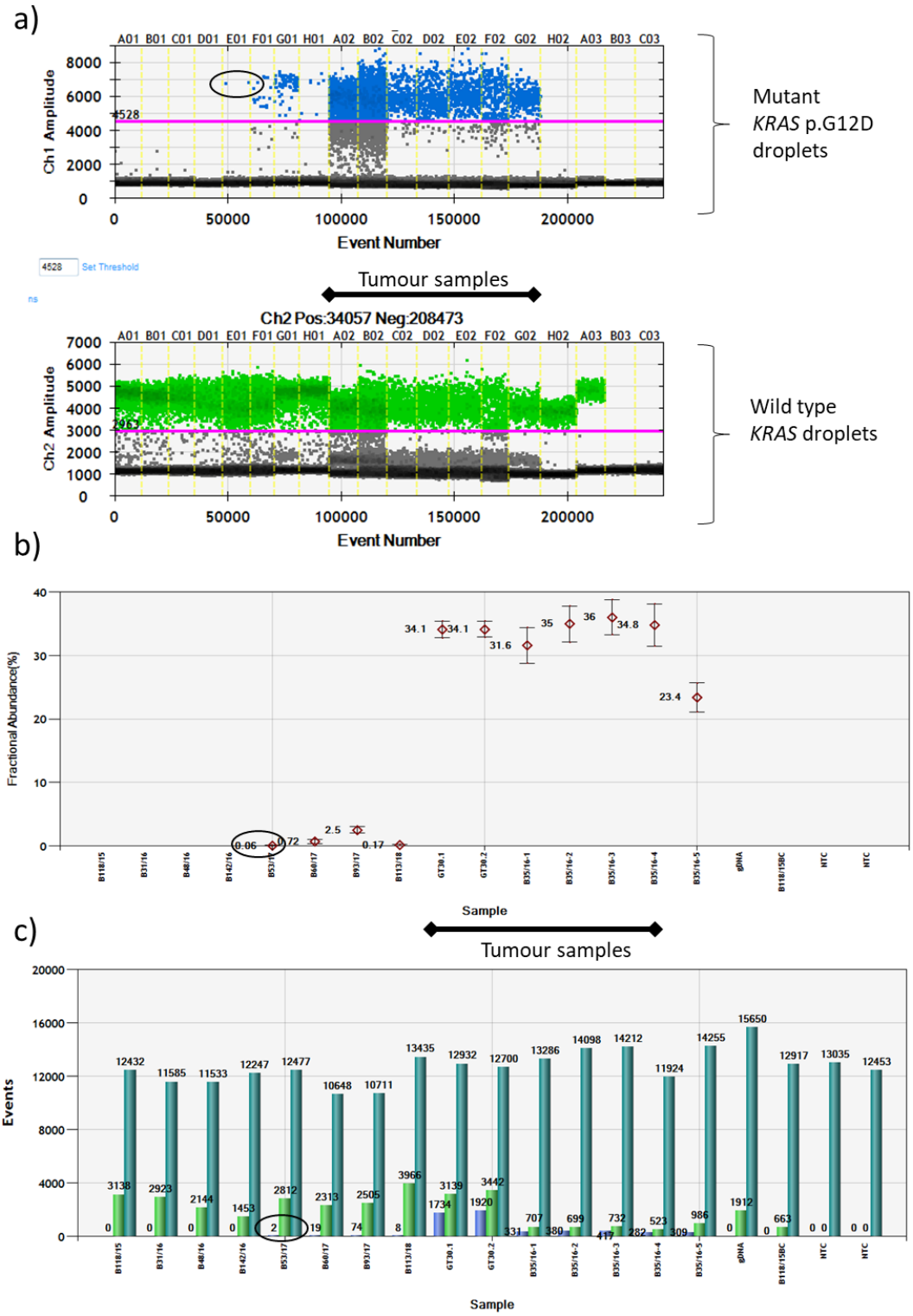
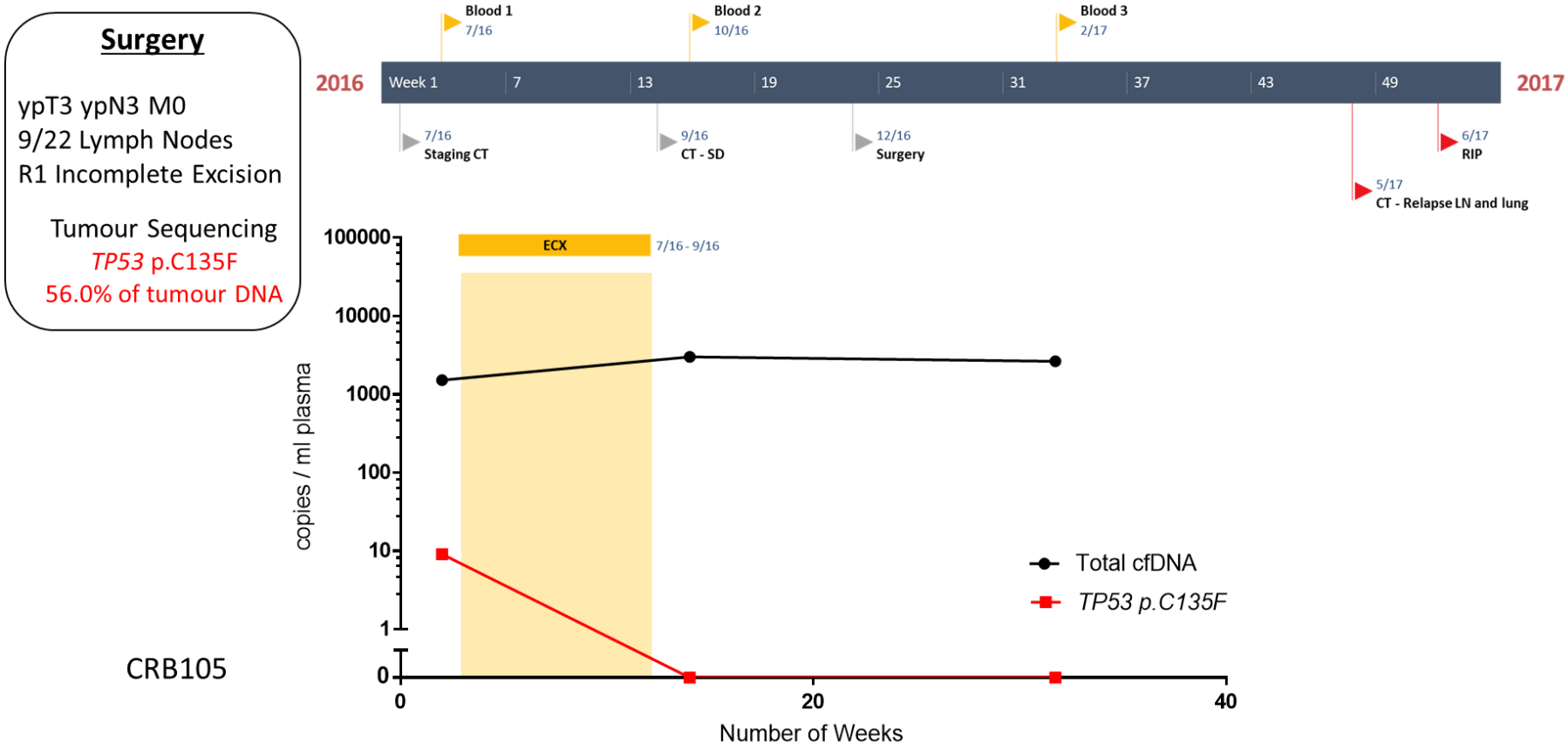


Figure 58 - Patient CRB58, ddPCR results for total cfDNA, tumour and buffy coat samples. a) Showing the raw data for presence of the KRAS p.G12D mutation b) Fractional abundance = VAF c) The total number of events (grey bar), and positive droplets for p.G12D mutant (blue bars) and wildtype (green bars) DNA. GT30.1+2 = OGD biopsy , B35/16 = Fresh frozen tissue samples 1-5, B118/15BC = buffy coat sample for CRB58, gDNA = genomic DNA, NTC = non-template control

Two other patients had a positive pre-neoadjuvant chemotherapy blood test, which resolved on treatment. Both patients went on to relapse but this was not detected by total cfDNA analysis.

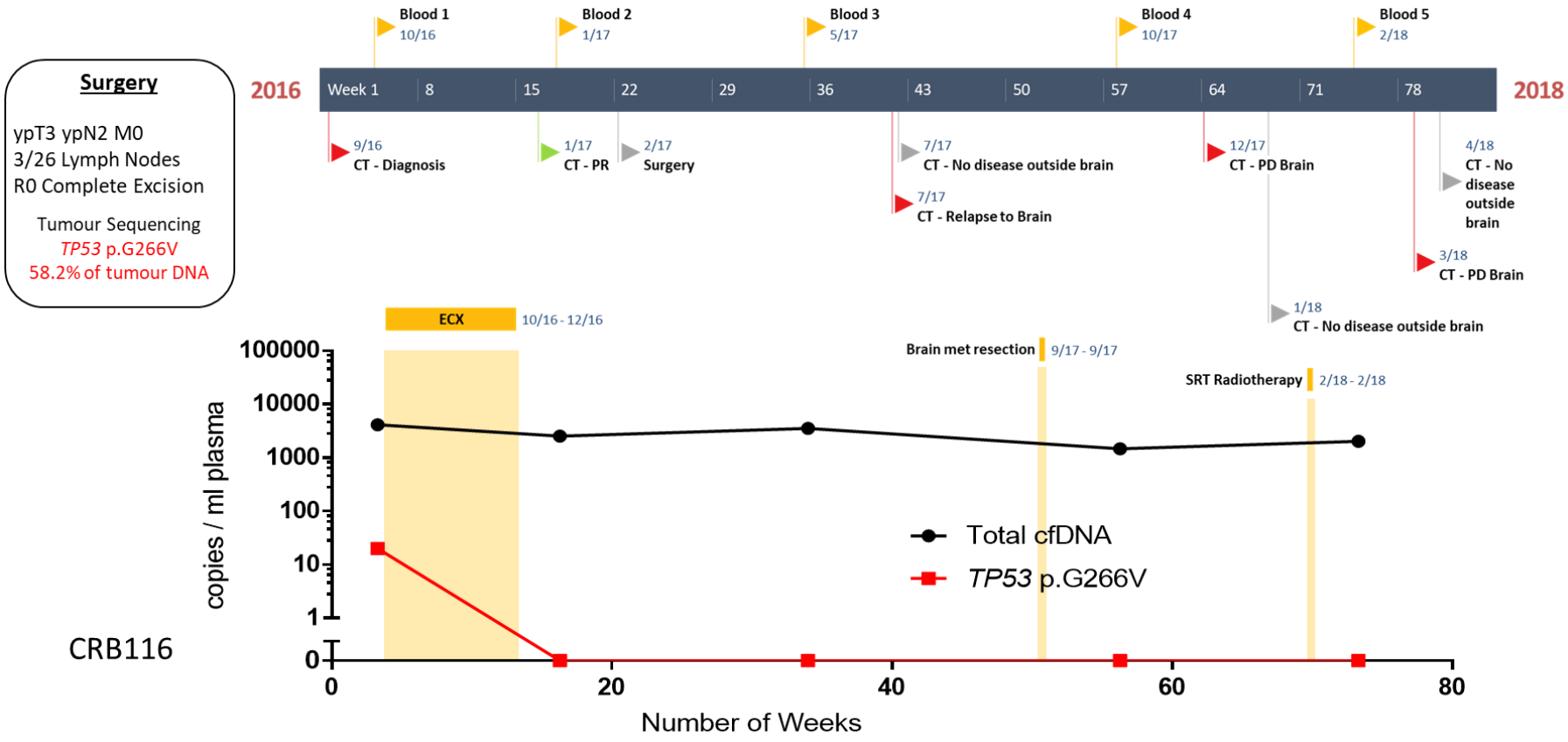
Patient CRB105 had a *TP53* p.C135F mutation (56.0% VAF) in the tumour DNA and which was also detected in the initial total cfDNA sample at 9 copies/ml (0.5% VAF). The mutation was not detected in the post-surgical sample, and the patient relapsed 15 weeks later (Figure 59). A blood sample was not available at relapse and would have been predicted to have been positive.

Patient CRB116 had a *TP53* p.G266V mutation (58.2% VAF) in the tumour DNA. This mutation occurred one base from the end of the GOA *TP53* amplicon 2 (Table 25) and was therefore the only mutation not to meet the guidelines set out in section 2.13.4. However, it was confirmed by a specific ddPCR assay and therefore cfDNA samples were screened. Analysis of total cfDNA showed detection of the *TP53* p.G266V mutation in the initial blood sample only (Figure 60). All the other samples were negative including those taken after metastatic relapse. Of note the patient relapsed to brain only, with no evidence of systemic disease. For this reason, the patient underwent resection of their brain metastasis and radiotherapy following local relapse to the same brain site.



Blood number		1	2	3
cfDNA copies /ml		1515	3000	2640
TP53 p.C135F	AF (%)	0.5	0	0
	Mutant copes / ml	9	0	0

Figure 59 - **Timeline for patient CRB105.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.C135F mutation detected in the total cfDNA by ddPCR also in copies/ml (red). Note the total cfDNA was only detectable before surgery. SD = stable disease, LN = Lymph node.



Blood number		1	2	3	4	5
cfDNA copies /ml		4141	2525	3535	1470	2030
KRAS p.G12D	AF (%)	0.48	0	0	0	0
	Mutant copies/ml	20	0	0	0	0

Figure 60 - **Timeline for patient CRB116.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.G266V mutation detected in the total cfDNA by ddPCR also in copies/ml (red). Note the total cfDNA was only detectable before surgery. SD = stable disease, LN = Lymph node.

5.6.2.3 Patients who relapsed without detectable ctDNA

Two patients who relapsed had no evidence of ctDNA at any time point based on tracking the main mutation detected in the primary tumour using the ampliseq GOA panel.

Patient CRB71 (Figure 61), had a high frequency *TP53* p.R248Q mutation (35.3% VAF) in the tumour. There was no evidence of this mutation by ddPCR in all 3 blood samples prior to relapse (bloods 1-3) and in the single blood sample (blood 4) following relapse to local lymph nodes. Interestingly CRB71 remains alive at the last review in May 2018, 12 weeks after relapse despite declining further medical intervention. Total cfDNA levels were low overall at 700-2200 copies per/ml

Patient CRB80 had no detectable ctDNA, and a *TP53* p.R175H mutation (6.4% VAF) in the resected primary tumour DNA (Figure 62). This patient was unusual in that relapse occurred to the muscle within the leg. It was possible that the dominant mutation had not been identified at this frequency, and as a result both the initial OGD biopsy (pre-resection) and relapse tumour within the leg were tested by ddPCR for the same mutation. The initial OGD biopsy showed the same mutation at 44.9% VAF indicating that this was likely a high frequency truncal mutation. Its low frequency in the surgical sample may be due to response to chemotherapy, or a low percentage of tumour cells in the sample. The sample following surgical excision of the tumour in the leg had the same *TP53* p.R175H mutation at a frequency of 27.5% indicating that this tumour is indeed a metastasis from the original GOA. Total cfDNA levels remained low through 4 serial blood samples at 900-1700 copies/ml.

CRB71

Blood number		1	2	3	4
cfDNA copies /ml		2272	985	1818	767
TP53 p.R248Q	AF (%)	0	0	0	0
	Mutant copies/ml	0	0	0	0

Figure 61 - **Timeline for patient CRB71.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.R248Q mutation not detected in the total cfDNA by ddPCR (red). PR = Partial response, LN = Lymph node.

CRB 80

Blood Number		1	2	3	4
cfDNA copies /ml		1689	1280	901	1737
TP53 p.R175H	AF (%)	0	0	0	0
	Mutant copies/ml	0	0	0	0

Figure 62 - **Timeline for patient CRB80.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.R175H mutation not detected in the total cfDNA by ddPCR (red).

5.6.2.4 Patients who did not relapse

By the arbitrary census date of 4th May 2018, 11 patients with a trackable mutation in tumour DNA had not relapsed. 10 of these 11 patients were negative for ctDNA in all blood samples (Appendix 7.8). One patient only had a positive ctDNA sample; the post-operative blood sample for CRB137 and the subsequent blood sample for the same patient was negative (Figure 63).

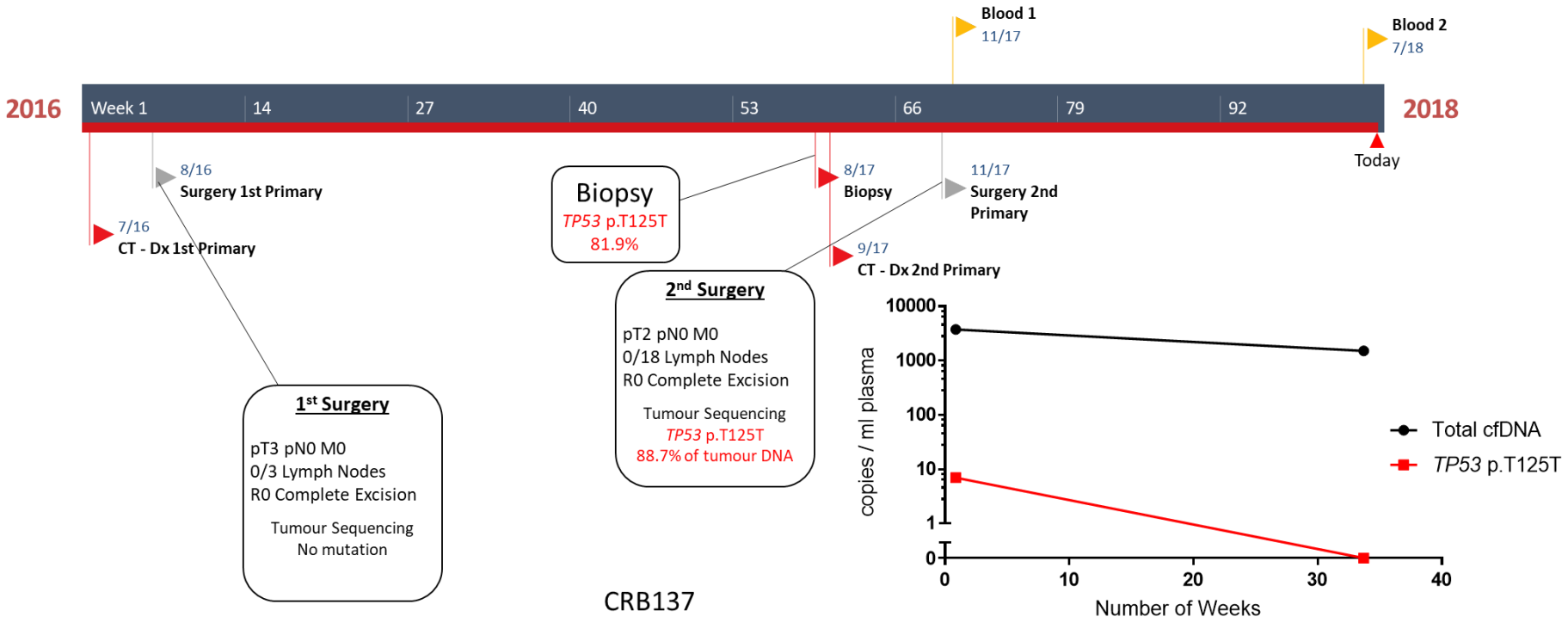
The patient was alive and relapse free 34 weeks (237 days) following surgery. The patient had two separate surgeries for two primary gastric tumours, which according to their clinical characteristics had been classified as unrelated; meaning two separate primary tumours had occurred sequentially for this patient.

To confirm whether the two cancers were related, both the primary tumours were sequenced. The 1st primary sequenced very poorly but appeared to reveal a *KRAS* p.V8V mutation with a 98% VAF. The second tumour showed a *TP53* p.T125T (88.7% VAF) and the same *KRAS* p.V8V mutation at 37.4% (Table 42). Literature review showed *KRAS* p.V8V was a silent mutation (COSMIC ID:COSM507) (Harlé et al., 2016), whilst the *TP53* p.T125T mutation, despite also being synonymous, was a pathological splice site mutation, at the final base of exon 3, resulting in dysfunctional TP53 protein (COSMIC ID:1638003) (Varley et al., 2001, Saiki et al., 2015). Sequencing of the buffy coat (lymphocyte) DNA, showed that the *KRAS* p.V8V mutation was an inherited mutation (Table 42) with a VAF of 48.8%, and was therefore likely an inherited SNP. The *TP53* p.T125T mutation was absent from buffy coat (germline) DNA suggesting it was a somatic mutation unique to the 2nd primary tumour. A ddPCR assay was designed for the *TP53* p.T125T mutation, allowing the tumours to tested using a second method (Figure 64). The results of this ddPCR, confirmed that the *TP53* p.T125T mutation was absent from the first primary, but present in both the biopsy and surgical resection of the second primary. This showed that the initial pathological finding was correct and this was a case of two primary gastric tumours.

The first total cfDNA sample tested positive for the *TP53* p.T125T mutation confirming presence of ctDNA. The subsequent blood sample taken at 33 weeks (230 days) following surgery did not detect ctDNA and the patient remains well

Run summary matrix					SNV specific matrix					
Sample	Mapped Reads	On Target %	Mean Depth	Uniformity %	Gene	Mutation	Coverage	Mutant Call	Freq %	Quality
1st Primary Tumour	177667	6.4%	89.37	77.49%	TP53	T125T	58	0	0	0
					KRAS	V8V	55	51	93	735
2nd Primary Tumour	567259	95.17%	13071	95.22%	TP53	T125T	24704	21911	88.8	81339
					KRAS	V8V	9661	3581	37.4	21194
Buffy coat	436996	98.74%	10681	95.06%	TP53	T125T	14221	1	0	0
					KRAS	V8V	5412	3137	48.8	26639

Table 42 - **Sequencing summary statistics, for tumours and buffy coat for patient CRB137**. Note poor on target call for 1st primary highlighted



Blood Number		1	2
cfDNA copies /ml		3730	1515
TP53 p.T125T	AF (%)	0.18	0
	Mutant copies/ml	7	0

Figure 63 – **Timeline for patient CRB137**. Showing the results of tumour sequencing and *TP53* p.T125T ddPCR of the biopsy and total cfDNA samples.

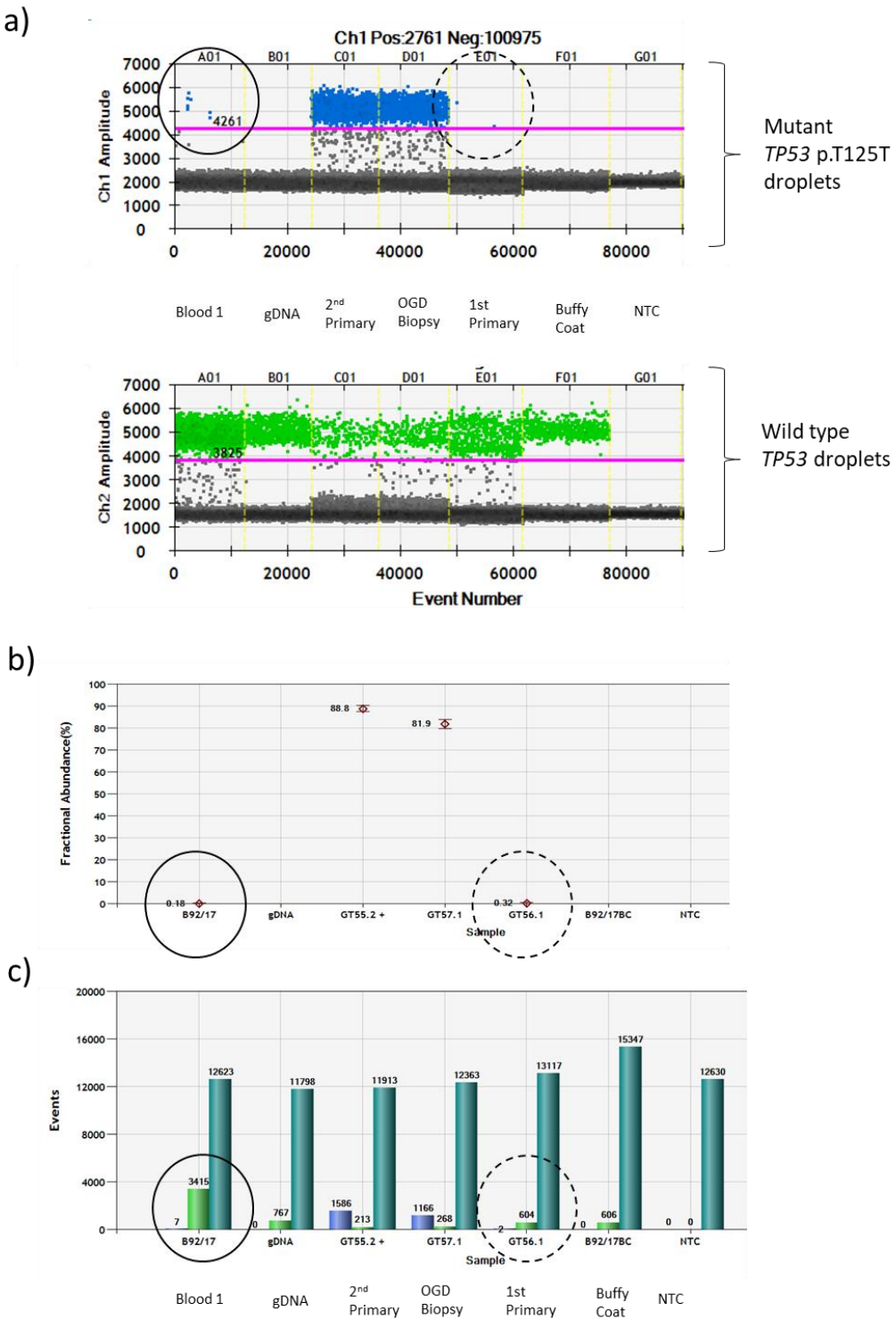
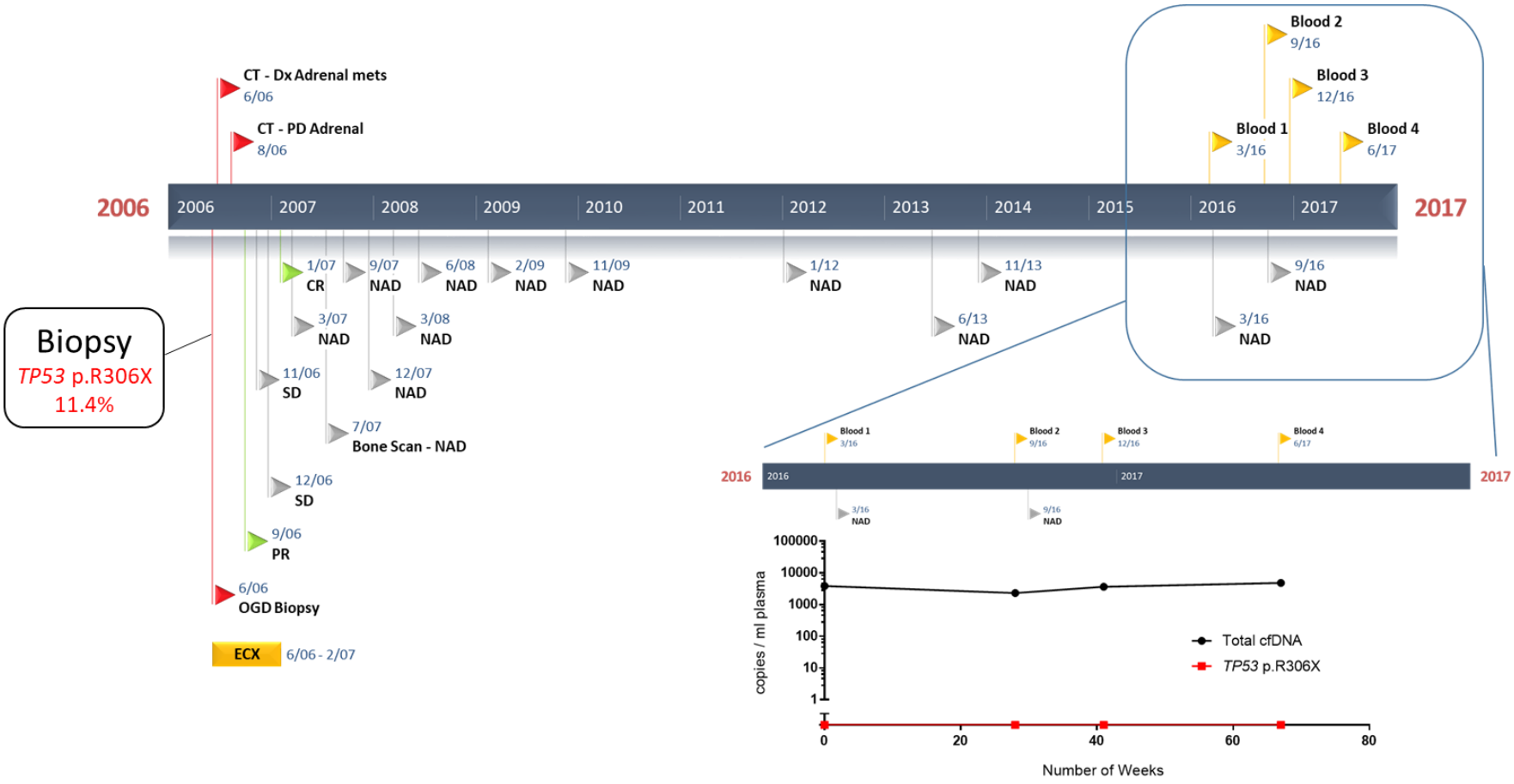


Figure 64 – **ddPCR result for CRB127 *TP53* p.T125T.** a) Raw ddPCR results b) Fractional abundance = VAF c) The total number of events, and positive droplets for mutant and wildtype DNA. Solid oval highlighting results for positive post-operative blood B92/17, dashed oval highlighting negative result from 1st primary tumour. gDNA = genomic DNA, GT55.2 = 2nd Primary tumour, GT57.1 = OGD biopsy of 2nd primary tumour, GT56.1 = 1st primary tumour, B92/17BC = buffy coat, NTC = non-template control.

5.6.3 Analysis of exceptional responder patients

Four patients were classified as exceptional responders, 1 of whom did not have tumour DNA of sufficient quality for sequencing. Two of the three (66%) other patients had no SNV detected in tumour DNA, and 1 patient had a detectable *TP53* mutation.

Patient CRB81 had an OGD biopsy that tested positive for *TP53* p.R306X (11.4% VAF) (Figure 65). The patient was initially diagnosed with metastatic GOA to the adrenal glands in 2006. Following disease progression, chemotherapy was given, which resulted in a partial response. A few months following completion of chemotherapy a complete response was noted. Given there was no chemotherapy given at this time, the cause of the complete response is unclear. This type of response is sometimes seen in other cancers treated with immunotherapy and it is possible that the immune system is responsible for the delayed complete response in this case. The patient had no detectable ctDNA as determined by ddPCR analysis of multiple blood samples for this mutation.



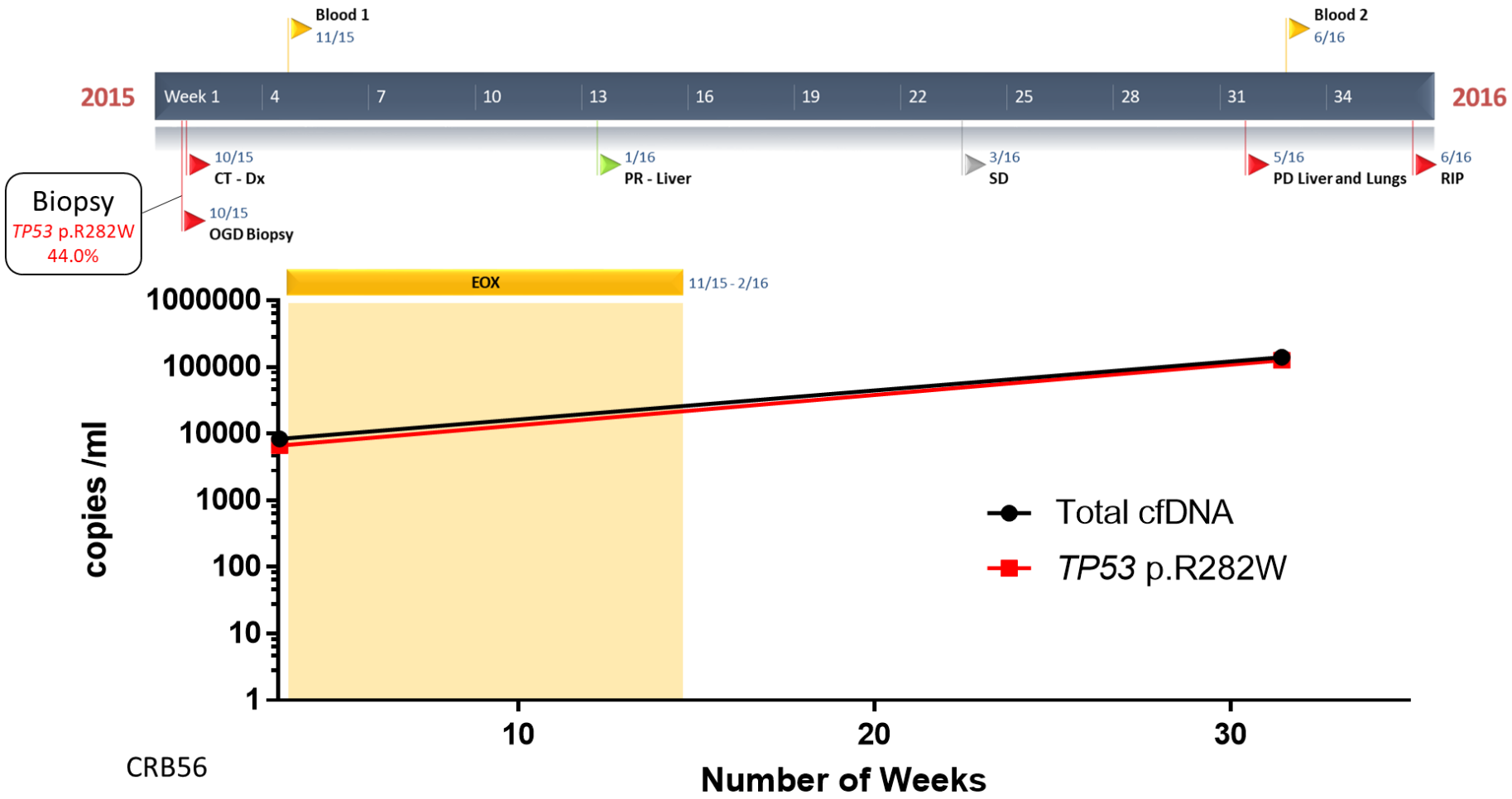
Blood number		1	2	3	4
cfDNA copies /ml		3863	2323	3636	4848
TP53 p.R306X	AF (%)	0	0	0	0
	Mutant copies/ml	0	0	0	0

Figure 65 - **Timeline for patient CRB81.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.R306X mutation, absent by ddPCR (red). PD = Progressive disease, PR = partial response, PD = progressive disease, SD = Stable disease. Lower timeline from includes period for which total cfDNA samples are available.

5.6.4 Case report of a patient with metastatic squamous cell oesophageal carcinoma

Patient CRB56 was the second patient recruited and was subsequently shown to have a squamous carcinoma rather than an adenocarcinoma. Sequencing of the tumour detected a *TP53* p.R282W mutation (44.0% VAF) for which a ddPCR assay was already available, which was used to analyse the two bloods available.

The first blood sample, taken before commencement on palliative chemotherapy had high levels of ctDNA (6,673 copies/ml), which increased significantly on progression of disease to liver and lungs (126,330 copies / ml) (Figure 66). The results suggest that squamous oesophageal cancer generates more ctDNA than GOA, as shown by the high VAFs of 79.5% and 89.5% respectively. A similar pattern of ctDNA positivity has been seen between squamous lung cancer and adenocarcinoma of the lung in the TRACERx trial (Abbosh et al., 2017).



Blood Number		1	2
cfDNA copies /ml		8393	141151
TP53 p.R282W	AF (%)	79.5	89.5
	Mutant copies/ml	6673	126330

Figure 66 - **Timeline for patient CRB56.** Showing clinical events in relation to total cfDNA copies/ml (black) and ctDNA as defined by detection of *TP53* p.R282W (by ddPCR) shown as copies/ml (red). SD = Stable disease, PD = progressive disease, PR = partial response.

5.6.5 Overview of results: SNV Analysis

5.6.5.1 Sequencing panel and sensitivity

The designed sequencing panel was effective at picking up high frequency mutations in the majority of tumour samples. However, it could not reliably detect mutations in the cfDNA below approximately 2% VAF. One reason for this was identified as the selection of slightly larger *TP53* amplicons (85bp vs 114bp) from the Acquires panel. Therefore, cfDNA samples that were negative via NGS panel sequencing were also analysed by ddPCR, and all borderline ddPCR samples were repeated with increased DNA input where sufficient DNA was available. 15 of 109 (13.8%) cfDNA samples analysed in this chapter had a tumour specific VAF >2%, whilst overall 29 of 109 samples (26.6%) had a detectable mutation VAF >0.05%, showing the importance of having used high sensitivity ddPCR to detect ctDNA.

5.6.5.2 Disease monitoring and lag time in metastatic patients

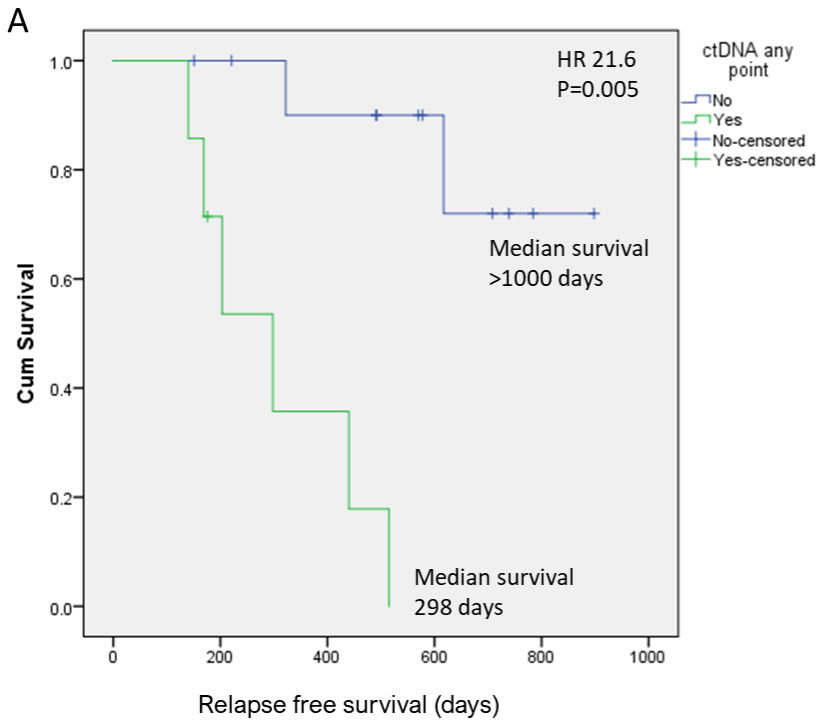
The timelines of those patients who had metastatic disease or had relapsed (sections 5.6.1.1, and 5.6.2.2) show that ctDNA levels are dynamic and fluctuate with treatment outcomes. There were a number of patients in whom ctDNA levels failed to fall during chemotherapy, who then relapsed relatively quickly, despite partial response on the relevant CT scan.

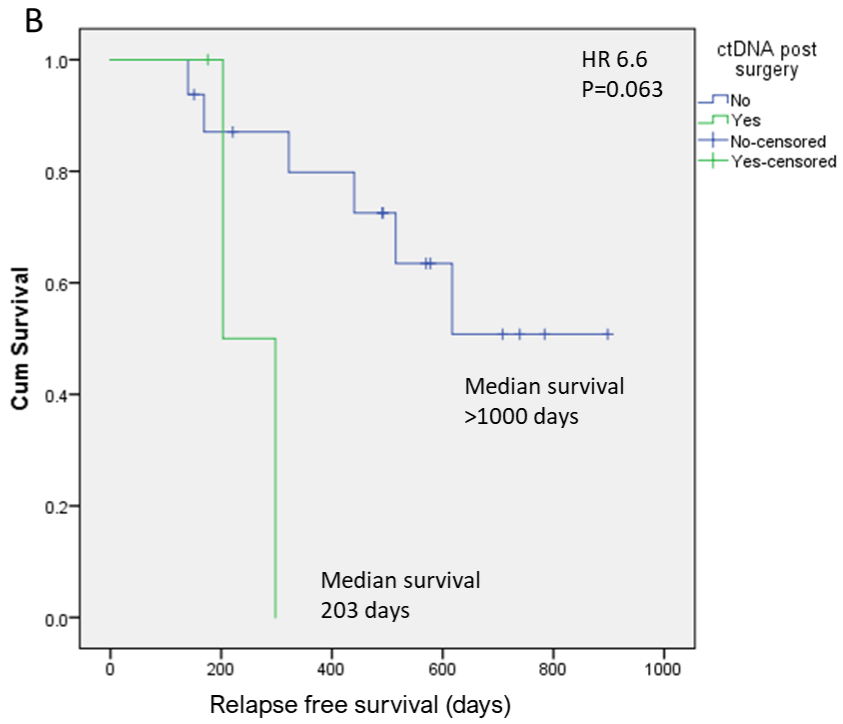
5.6.5.3 Survival and the presence of ctDNA

To investigate the prognostic significance of ctDNA, relapse free survival in patients treated with curative intent was investigated in those patients with and without detectable ctDNA. When the presence of ctDNA was considered in the post-surgical blood only (Figure 67a), there was a non-significant difference (Cox regression p=0.063) between those with and without detectable ctDNA. The lack of significance was likely due to the low numbers of patients with detectable ctDNA in the post-surgical blood (3). A fourth patient, CRB107 relapsed early before a post-surgical sample was taken. Had a post-surgical, pre-relapse

blood been present for this patient, it is likely that this would also have been positive – as both the preceding and proceeding blood tests were positive (Figure 56, page 197). It is important to note that had a further relapsed patient with positive ctDNA been present in this analysis, this would have resulted in a positive statistical test. These results also show that up to 20% of patients (4 out of 19 patients) undergoing surgery may have ctDNA positive post-surgical ctDNA bloods, which is of importance in designing future clinical trials using ctDNA.

When relapse free survival was plotted against presence of ctDNA at any time point pre or post surgery (Figure 67b), but before relapse, there was a significant difference between the two populations (Cox regression p=0.005, HR 21.6, CI 2.5-185.5) confirming that ctDNA was also a poor prognostic sign in the curative setting. This suggests that a patient who has detectable ctDNA at any time point before or after surgery is at high risk of relapse and this could include up to 35% (7 out of 19 patients in this study) of patients who successfully undergo surgery.





Days	0	200	400	600	800	1000
ctDNA Positive	7	4	2	0	0	0
ctDNA Negative	11	9	8	5	1	0

Days	0	200	400	600	800	1000
ctDNA Positive	3	2	0	0	0	0
ctDNA Negative	16	13	10	5	2	0

Variables in the Equation

	B	SE	Wald	df	Sig.	HR	95.0% CI for HR Lower	Upper
plasma ctDNA any point	3.077	1.095	7.896	1	.005	21.691	2.536	185.501

Variables in the Equation

	B	SE	Wald	df	Sig.	HR	95.0% CI for Exp(B) Lower	Upper
plasma ctDNA post surgery	1.890	1.016	3.459	1	.063	6.617	.903	48.471

Figure 67 – **Relapse-free survival, in time from surgery.** a) Presence of post- surgical ctDNA. b) presence of ctDNA at any timepoint pre-relapse. Number at risk shown below each graph. Cox regression analysis shown in table. N=19.

5.6.5.4 PEACE

Two patients have been recruited to the PEACE (Posthumous Evaluation of Advanced Cancer Environment) study. This study allowed sampling of patient's tumours after death via post mortem. Patient's CRB58 and 69 were recruited and post mortem dissection of tumour sites, has been undertaken.

For CRB58 the sites of tumours comprise (Figure 68);

1. An obstructive point in the biliary tree
2. A metastasis from the liver
3. Pleural effusion; liquid and pellet of cells.
4. A metastasis in the pontine region of the brainstem

Preliminary analysis of the pleural fluid sample has shown detection of the *KRAS* p.G12D mutation with a VAF of 5%, and the cellular pellet from the effusion had a VAF of 4%, indicating that the effusion was directly cancer related and not due solely to another pathological process. Also of note the VAF for the pleural effusion was higher than any of the total cfDNA samples (0.06-2.5%) from the blood (see Figure 57, page 199).

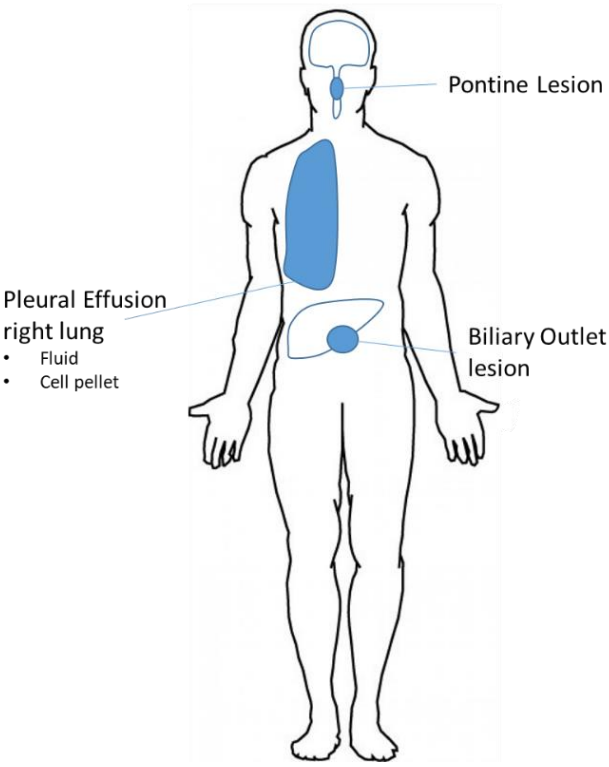


Figure 68 **– Patient CRB58: Sites of tumours identified and sampled at autopsy**.

For CRB69 the following tumours were sampled (Figure 69), but no analysis has yet been undertaken;

- 3 Brain Lesions
    - Right Superior parietal lesion
    - Right occipital Lesion
    - Left Cerebellar Lesion
- Primary GOJ Lesion
- Affected lymph nodes
    - Right Hilar node
    - Right Para-tracheal Station 4 node

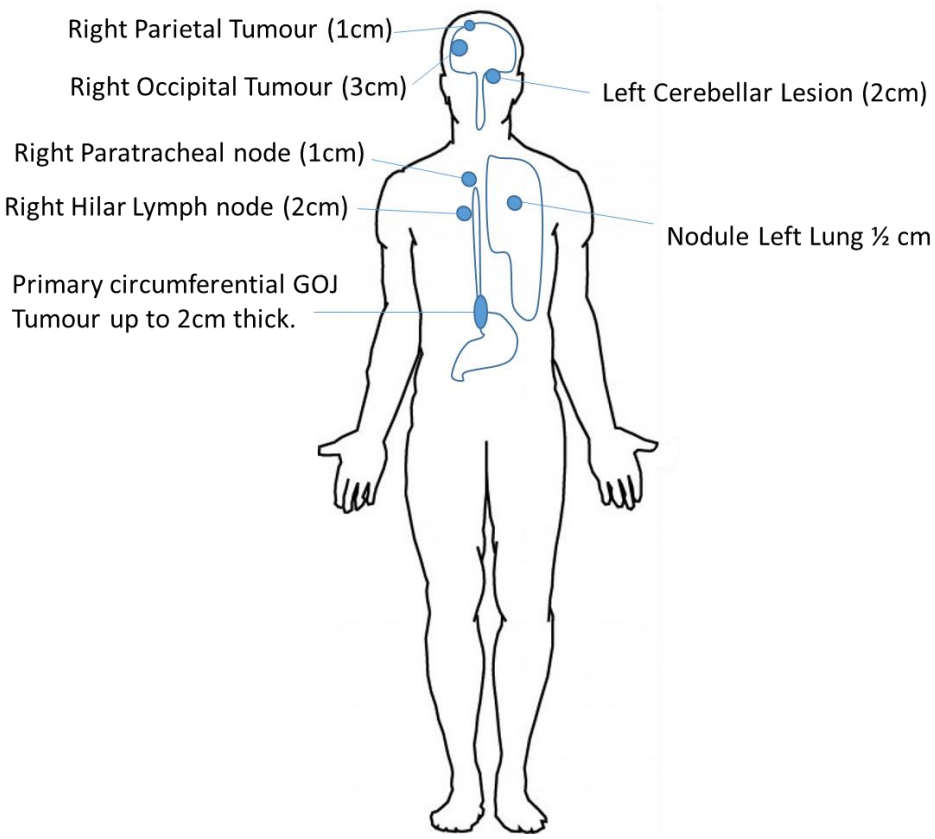


Figure 69 - **Patient CRB69: Sites of tumours identified and sampled at autopsy**.

5.7 Results: SCNA Analysis

Analysis of SCNA in the published genomic data has shown that GOA has significant levels of gene amplification (section 1.3.3.3). Therefore, it was particularly interesting to discover if the 5 metastatic or curative intent patients that underwent sequencing and had no detected SNV >5% VAF in their tumour DNA (CRB 78, 86, 94, 102, 109) had a detectable SCNA. Funding also allowed a further 15 patients to have analysis for SCNA, and therefore 20 patient tumour DNA samples were sent for NanoString analysis.

5.7.1 NanoString analysis of Tumour DNA

NanoString detects gene amplification using hybridisation probes (section 2.15 page 76). There are three probes for each gene to improve accuracy of copy number assessment. There is inherent variance in probe binding between samples due to the level of degradation of the DNA and the amount of DNA added to the reaction. Mathematical calculation allows correction due to difference in probe binding due to these two issues (section 2.15.2) but these variances cannot be removed completely. To reduce false positive calling of SCNAs an amplified gene was defined as;

1. Amplification in copy number across all three probes in the gene
2. Copy number of equal to or greater than 4

In total 20 patients had 400ng of tumour DNA analysed by NanoString nCounter, including all 5 of the patients who had no mutation following sequencing. A selection of FFPE and Fresh frozen derived DNA samples were sent. DNA from fresh frozen tissue showed less variance in counts between samples and could be compared to a standard control. FFPE tissue showed more variance between samples, and where possible the control used was normal tissue from the same surgical or biopsy sample. If normal tissue was unavailable two different control FFPE samples were used to reduce variance caused by control tissue selection. Of the 20 patients analysed, 13 (65%) had one or more detectable SCNA (Figure

70). The majority of genes were amplified with a copy number of less than 10, but in three cases it was greater than 30.

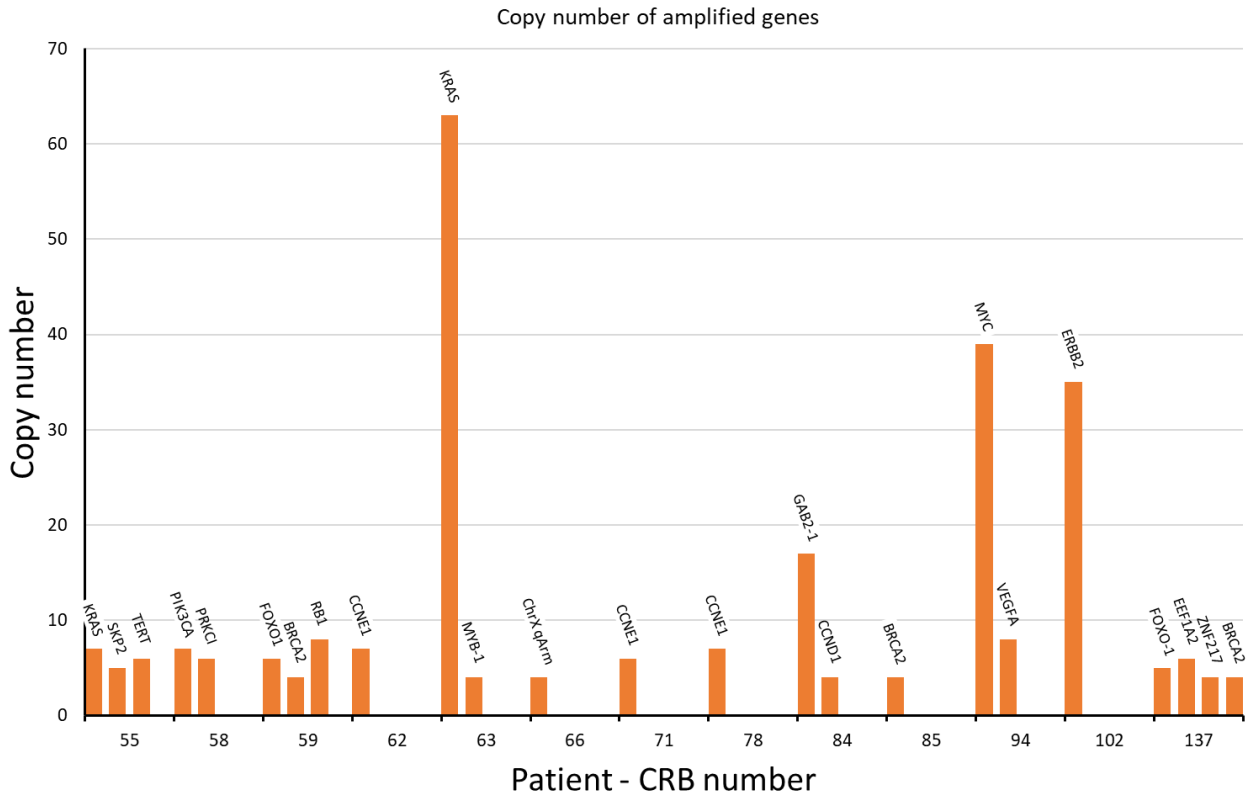


Figure 70 - **Copy number of amplified genes, from tumour DNA samples sent to nanostring**

Of the 5 patients with an undetectable SNV that were sent to NanoString for analysis, 3 had an identified SCNA, and one had two amplified genes as shown below (Table 43).

Patient	Cohort	Gene	Copy number	Gene	Copy number
CRB78	Curative	*CCNE1*	7		
CRB86	Palliative	None			
CRB94	Curative	*MYC*	38	*VEGFA*	8
CRB102	Curative	*ERBB2*	35		
CRB109	Curative	None			

Table 43 – **SCNAs in patients with no detected SNV**

5.7.2 Confirmation of copy number by ddPCR and qPCR

The 3 patients with no SNV detected by sequencing and SCNA detected by NanoString nCounter were of particular interest to see if ctDNA could be detected in these patients. Initially the highest frequency SCNAs in these patients were confirmed by ddPCR and qPCR.

Analysis via qPCR was carried out using three separate standard gene primer probe assays against *RPPH1, GAPDH* and *CNTNAP1 (*section *2.10.2)*. These assays had previously been designed and assessed for efficiency by the lab, and been shown to represent three different genes that are rarely amplified in cancer and could therefore be used as good reference genes (Page et al., 2011). Given the finding from analysis in CRB24 (Section 4.7) it was decided that a single reference gene should not be relied on when calculating copy number, as even these references can, albeit rarely, be amplified. Analysis via ddPCR for each sample was run twice using different two reference genes; *GAPDH* and *CNTNAP1*. The *RPPH1* reference was shown to be unreliable when used in ddPCR for cfDNA analysis.

These experiments show that both ddPCR and qPCR results broadly agree with the results of NanoString nCounter, which tends to slightly underestimate the copy number (Figure 71).

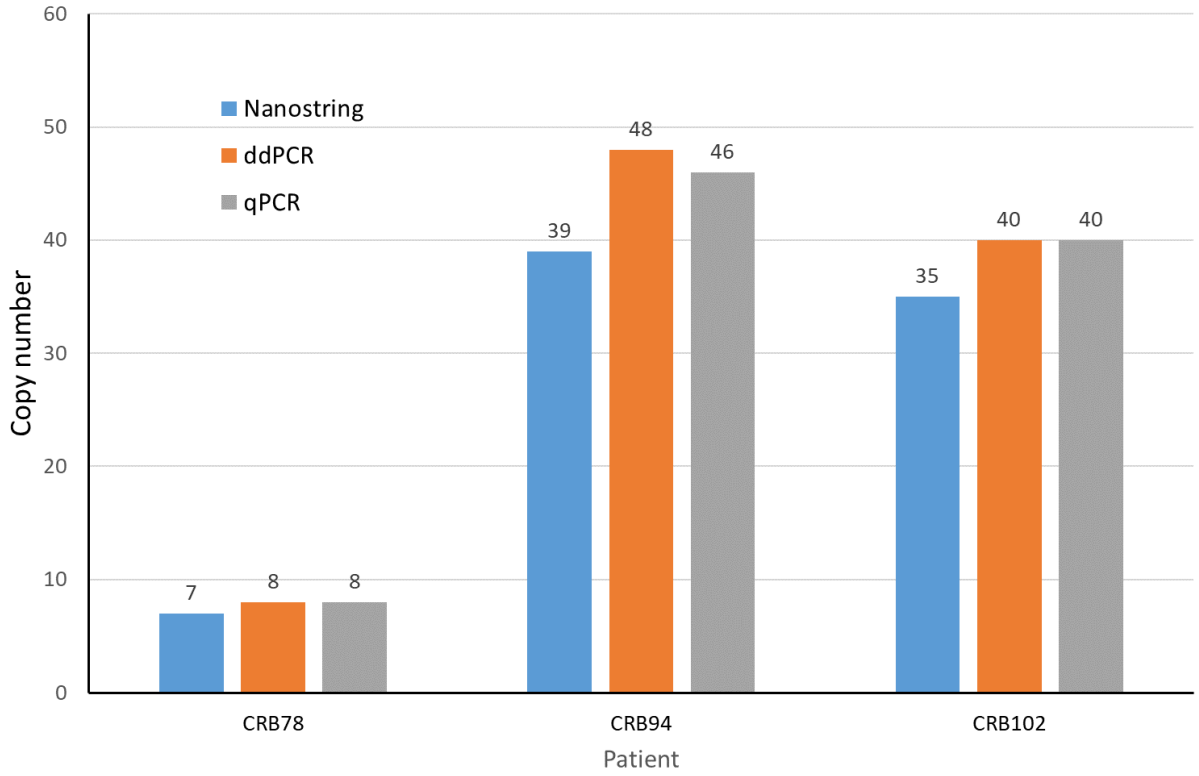


Figure 71 – **Copy number as determined by Nanostring, ddPCR and qPCR.**

5.7.3 SCNA Analysis of ctDNA

Given the positive findings of ctDNA detection using SNV analysis, the next experiment was to look for evidence of ctDNA in the three patients proven to have a SCNA. The specific aim was to look for evidence of minimal residual disease. Also by chance two of these patients (CRB94 and 102) had high copy numbers, > 30 CN. The total cfDNA samples were analysed by both ddPCR and qPCR (Table 44).

Patient	Tumour Gene	CN	Blood	ddPCR *GAPDH*	*CNTNAP1*	Average	qPCR *RPPH1*	*GAPDH*	*CNTNAP1*	Average	Overall Average
94	MYC	45	1		fail		1.8	2.8	1.67	2.09	2.09
			2	2.53	1.8	2.17	2.07	2.82	1.86	2.25	2.21
			3	3.7	2.25	2.98	2.72	3.54	2.13	2.80	2.89
78	CCNE1	8	1	2.83	1.58	2.21	2.73	3.52	2.086	2.78	2.49
			2	2.43	1.54	1.99	1.9	2.74	1.74	2.13	2.06
			3	fail	1.6	1.60	2.85	2.73	1.8	2.46	2.03
102	ERBB2	40	1	3.65	2.2	2.93	2.79	3.37	2.14	2.77	2.85
			2		fail		3.52	3.23	2.11	2.95	2.95
			3	3.24	2.1	2.67	2.84	2.69	1.88	2.47	2.57
			4	2.88	2.12	2.50	3.07	3.18	2.2	2.82	2.66
			5	3.2	2.02	2.61	2.99	2.96	2.13	2.69	2.65

Table 44 – **Comparative results of ddPCR and qPCR analysis of total cfDNA from CRB94, 78 and 102.** The relapse blood sample for patient CRB94 is highlighted.

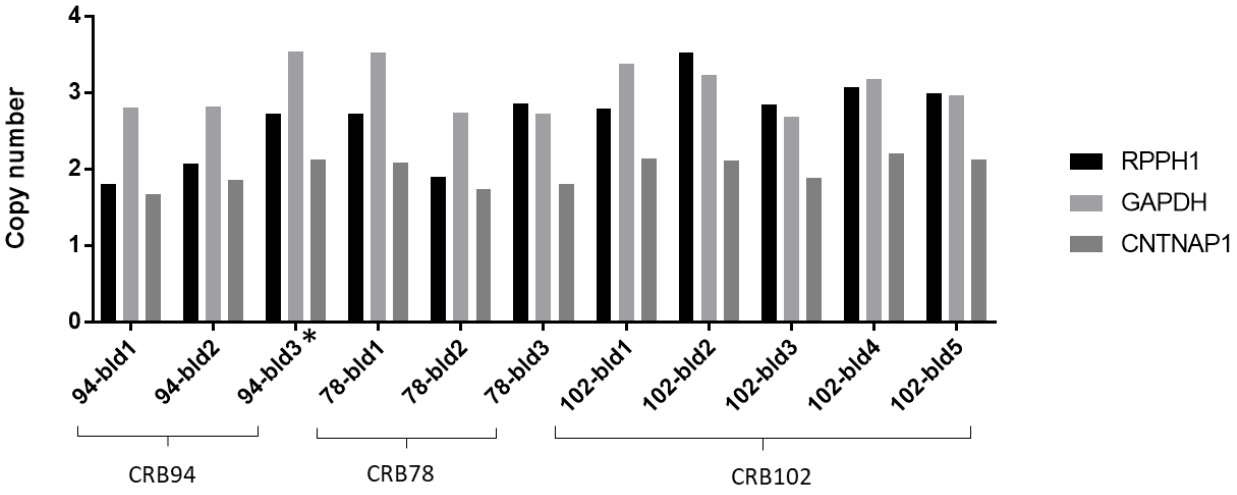


Figure 72 – **Analysis of tumour specific gene amplification in total cfDNA by qPCR.** CRB9 =*MYC* amplification, CRB78=*CCNE1* amplification, CRB102=*ERBB2* amplification. *Relapse blood of Patient CBR94.

Applying a threshold CN=3 for amplification, and using average values calculated using either ddPCR or qPCR results, no sample showed evidence of copy number gain in the plasma.

Of note the qPCR results were overall more reliable than ddPCR as there were no qPCR failures. The variability of the results was also quite large, with experiments using the reference gene *CNTANP1* recording the lowest copy number value and *GAPDH* the highest. The largest difference in calculated CN between reference genes was a CN of 1.41 via qPCR and 1.35 via ddPCR.

Patient CRB94 relapsed 7 days after blood 3 was taken. This blood test was therefore of importance in deciding if minimal residual disease could be detected by CN analysis. Following the results (Figure 72, Table 44) there is an upward trend in *MYC* copy number for CRB94 in the 3 consecutive total cfDNA samples, with an overall copy number average of 2.09, 2.21, 2.89 respectively. This rise in calculated CN could be consistent with disease recurrence, but when compared with the results from the bloods of those patients that did not relapse (CRB 78 and 102) the results for CRB94, and in particular blood 3, were not different to those of the other samples. This suggests that the use of this method to detect minimal residual disease is not possible.

5.8 Results: HER2 / *ERBB2* Analysis

Aim 3. of this project was to determine if *ERBB2* SCNA analysis of total cfDNA and tumour tissue can be used instead of traditional HER2 IHC to determine HER2 positivity.

Trastuzumab is the only targeted drug available in the treatment of gastric and GOJ cancers. IHC testing for HER2, followed by FISH analysis for patients that are recorded as HER 2+, is currently the gold standard method for determining if patients are HER2 positive, as it is for breast cancer patients. This determines if the patient will respond to the drug trastuzumab. Work by Shoda et al. (2017) has shown that HER2 status can be determined by ddPCR and qPCR in both tissue and total cfDNA, using the *ERBB2:RPPH1.*

Only 3 of the 44 original patients had metastatic disease that was HER2 positive. Therefore, an expansion cohort of 11 patients was identified from all the patients in the last two years that had received treatment with trastuzumab. This formed a HER2 expansions cohort of 14 patients for which tumour DNA was available. 6 of these 14 patients were alive and consented to blood sample donation for total cfDNA analysis.

5.8.1 *ERBB2* amplification in Tissue

Amplification of ERBB2 in tumour DNA was detected by ddPCR. As the *RPPH1* primers worked well for FFPE tissue is was possible to run all three reference genes (*GAPDH, RPPH1* and *CNTNAP1)*. A CN of 4.2 was used as the cut-off for amplification using the threshold previously determined by Page et al. (2011) for *ERBB2* in breast tumour tissue and plasma. The *GAPDH* reference provided the most discrimination of tumour samples. 8 of the 14 (57.1%) HER2 positive samples were amplified using a *ERBB2:GAPDH* ratio and 1 of the 8 (12.5%) HER negative samples was *ERBB2* amplified (Figure 73).. Highly *ERBB2* amplified samples (CN >30) were only present in IHC/FISH HER2 positive tumours (Figure 73). A ROC analysis of the data using the *ERBB2:GAPDH* tumour ratio (and a cut-off of CN=4.20) to discriminate between HER2 positive and HER2 negative tumours (as defined by IHC) had a sensitivity of 50.0% and specificity of 87.5%(Figure 74). The best discriminative power was

a CN of 3.17 with a sensitivity of 64.3% and sensitivity of 87.5%. The ROC analysis was however not statistically significant. The lack of a statistical difference is due in part to low sample size but also due to the fact that many HER2 positive patients (by IHC) do not appear to have tumour DNA amplified for *ERBB2*.

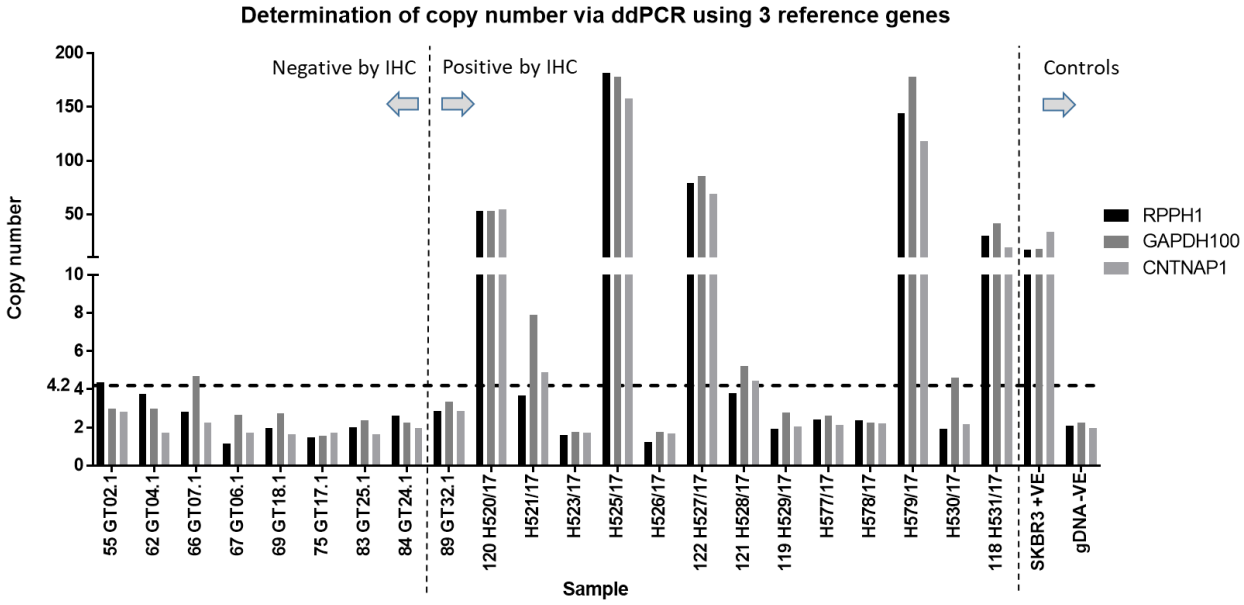
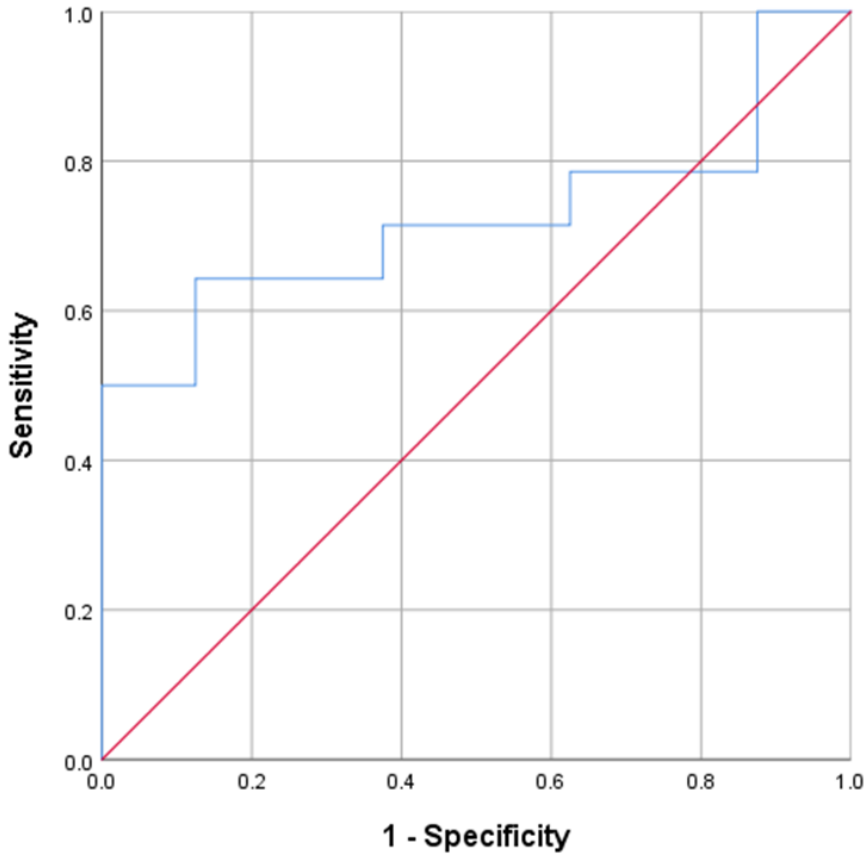


Figure 73 – **HER2 positive and negative tumours, analysed for *ERBB2* amplification by ddPCR.** Using three reference genes, *RPPH1, GAPDH* and *CNTNAP1*. Cut-off copy number value of 4.2 shown.



p-value = 0.088
AUC = 0.723
95% CI = 0.509-0.937

Figure 74 – ROC analysis of the data using the *ERBB2:GAPDH* ratio to discriminate between HER2 positive and HER2 negative tumours (as defined by IHC)

5.8.2 *ERBB2* amplification in total cfDNA

Six of the HER2 positive patients had total cfDNA samples available for analysis of *ERBB2* amplification. This was carried out via qPCR given the limited quantity of available total cfDNA and because of the increased reliability of this method in the previous section for total cfDNA analysis versus ddPCR.

Total cfDNA samples from the 6 HER2 positive patients were run alongside 7 samples from HER2 negative patients and from 2 healthy controls (Figure 75). The HER2 negative patients had the first blood analysed because Shoda et al. (2017) had shown that patients who are

initially *ERBB2* unamplified could acquire amplification over time and I wished to reduce the chance of this occurring. The same cut-off of CN=4.2 was used to determine amplification using *GAPDH* as the reference gene (Table 45). Using this cut-off 2 of the 6 (33.3%) IHC/FISH positive patients were negative by total cfDNA qPCR analysis, one of whom (CRB119) also had tumour DNA negative for amplification. Two of the 8 (25%) IHC negative patients were positive by total cfDNA qPCR analysis, including the single IHC negative patient that had *ERBB2* amplification in tumour DNA (CRB62). One patient, CRB118 had highly amplified total cfDNA.

ROC analysis showed that cfDNA analysis could discriminate HER2 positive from HER2 negative patients (p=0.022) (Figure 76) . With a cut-off of CN=4.20 the sensitivity was 66.7% and specificity was 80.0%. The best discriminative power was a CN of 3.96 with a sensitivity of 77.8% and sensitivity of 80.0%, which is reassuring close to the pre-defined CN of 4.20. If CRB119 was removed from the analysis because this was negative by both tissue and total cfDNA *ERBB2* analysis, and so by extrapolation would have been expected to be negative by FISH which specifically detects amplification, the ROC analysis was more highly significant (p=0.004).

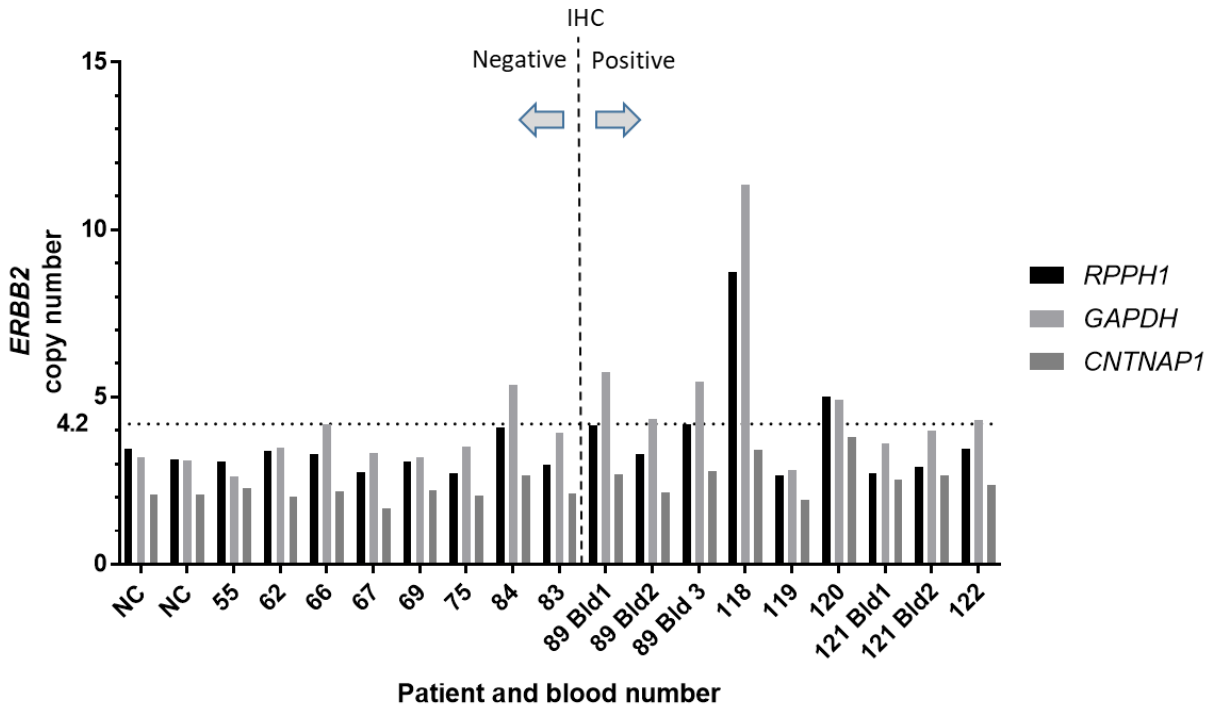
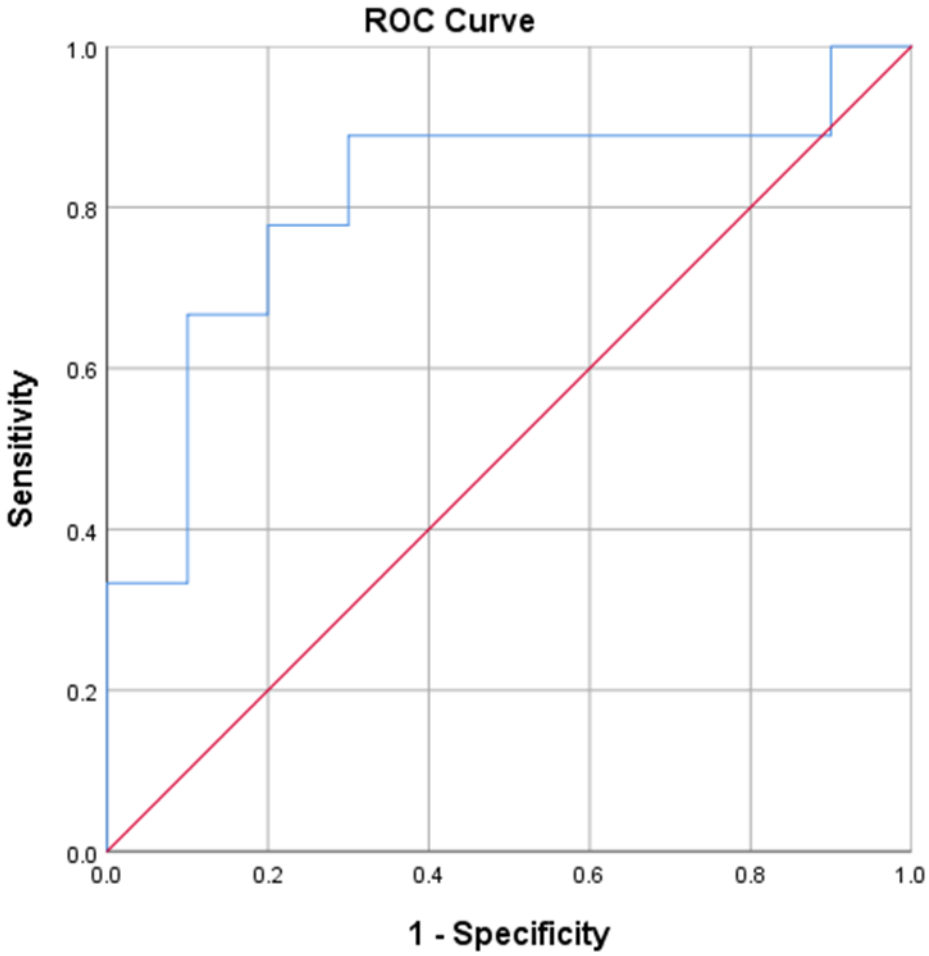


Figure 75 – **Summary barchart showing total cfDNA *ERBB2* copy number determined by qPCR for patients with HER2 negative and positive tumours by IHC.** Using three references, *RPPH1, GAPDH* and *CNTNAP1*. Cut-off copy number value of 4.2 shown. NC = normal control.

	Patient			Tissue by ddPCR		total cfDNA by qPCR		
	CRB	Tissue sample	IHC Result	Result	CNV	Blood	Result	CNV
IHC Negative	55	GT02.1	1+	NEGATIVE	2.99	1	NEGATIVE	2.62
	62	GT04.1	1+	NEGATIVE	2.98	1	NEGATIVE	3.48
	66	GT07.1	1+	POSITIVE	4.67	1	POSITIVE	4.20
	67	GT06.1	0	NEGATIVE	2.65	1	NEGATIVE	3.34
	69	GT18.1	2+ FISH neg	NEGATIVE	2.73	1	NEGATIVE	3.20
	75	GT17.1	0	NEGATIVE	1.57	1	NEGATIVE	3.51
	83	GT25.1	2+ FISH neg	NEGATIVE	2.39	1	NEGATIVE	3.93
	84	GT24.1	2+ FISH neg	NEGATIVE	2.25	1	POSITIVE	5.37
IHC Positive	89	GT32.1	3+	NEGATIVE	3.36	1	POSITIVE	5.73
						2	POSITIVE	4.33
						3	POSITIVE	5.46
	120	H520/17	3+	POSITIVE	53.00	1	POSITIVE	4.90
	122	H527/17	3+	POSITIVE	86.00	1	POSITIVE	4.32
	121	H528/17	2+ FISH pos	POSITIVE	5.20	1	NEGATIVE	3.60
						2	NEGATIVE	4.01
	119	H529/17	2+ FISH pos	NEGATIVE	2.79	1	NEGATIVE	2.82
	118	H531/17	3+	POSITIVE	41.80	1	POSITIVE	11.35

Table 45 – **Overview of HER2 IHC status and tumour DNA and total cfDNA *ERBB2* copy number as determined by *ERBB2:GAPDH*.** Copy number cut off for positively was 4.2.



p-value = 0.022
Area Under the Curve (AUC) = 0.811
95% Confidence Interval = 0.598-1.000

Figure 76 – ROC analysis of the data using the *ERBB2:GAPDH* ratio to discriminate between HER2 positive and HER2 negative patients as defined by tumour tissue IHC.

5.9 Discussion: GOA and ctDNA

Results in this chapter show that of 33 patients with a trackable mutation 16 had detectable ctDNA confirming the hypothesis that ctDNA is detectable using SNV analysis of ctDNA. The most commonly detected mutations were in *TP53* and the presence of these mutations was proven by two methods; ampliseq GOA panel sequencing and SNV targeted ddPCR, in both tumour DNA and total cfDNA.

In those patients with multiple blood samples and metastatic disease, ctDNA levels tracked disease response and progression. In metastatic disease high levels of ctDNA at diagnosis were shown to be a poor prognostic sign, and in patients treated with curative intent the presence of ctDNA at any time point before relapse was shown to be a significant indication of poor survival. In 3 patients the presence of ctDNA in the post-surgical blood indicated the possible presence of minimal residual disease and thus far 2 of these 3 patients have relapsed. ctDNA was absent in the single exceptional responder with GOA who had a trackable mutation but was detected in a single patient with squamous cell oesophageal cancer.

In the majority of patients relapse of disease was associated with presence of ctDNA. Of significance, 2 patients, CRB105 and 116, relapsed without detectable ctDNA in their post surgical blood but both patients had detectable ctDNA pre-surgery which disappeared with neoadjuvant chemotherapy. This suggests two points; firstly that ctDNA monitoring can detect the effect of neoadjuvant chemotherapy, and secondly that the presence of ctDNA is a significant detrimental finding, in keeping with previous results in other cancers (O'Leary et al., 2018, Diehl et al., 2008, Shoda et al., 2017, Lipson et al., 2014) . The results also suggest that even when ctDNA disappears, presumably due to the action of chemotherapy, disease still returns. Negative ctDNA, following prior positive ctDNA samples were also seen following neoadjuvant treatment in CRB61, adjuvant treatment in CRB55 and palliative treatment in patients CRB62, 69, and 75, but all patients had subsequent return of their disease. Therefore, disappearance of ctDNA can monitor chemo-responsive disease, but

does not indicate eradication of disease. It also suggests that detection of ctDNA at any time point is an indication of aggressive disease with poor survival.

5.9.1 Patient without detectable ctDNA and recurrent or metastatic disease

There were 6 patients who had relapsed or metastatic disease with trackable mutations but without detectable ctDNA in any blood test (CRB63, 64, 71, 74, 80, and 121); comprising 2 relapsed patients treated with curative intent and 4 metastatic patients. Patient CRB64 had equivocal ctDNA detection and so is not discussed here. There are a number of reasons why these patients may have had no detectable ctDNA:

1. The patient had ctDNA present, but the ctDNA remained undetected because
    a. The GOA panel did not detect the dominant mutation in the primary tumour or
    b. the mutation profile of metastatic disease had evolved since the primary tumour was sampled resulting in a different dominant mutation.

Patients CRB55 and CRB63 demonstrate these issues. Both had two high frequency mutations detected in tumour DNA. CRB55 had a *TP53* p.R196X and a *PIK3CA* p.I338N mutation in the primary tumour of which the *TP53 p.*R196X mutation was detectable in the total cfDNA but the *PIK3CA* mutation was not (Figure 54, page 194). This indicated loss of the *PIK3CA* p.I338N mutation in metastatic disease. CRB63, had a *TP53* p.R282W and a *CDK2NA* p.R80X mutation present in tumour DNA neither of which were detectable in total ctDNA. The *CDK2NA* mutation was initially missed by the GOA panel and only revealed by additional sequencing of both primary and metastatic sites with a larger panel (Table 39, page 184). This patient had high levels of the *TP53* p.R282W mutation in the metastatic tumour only, showing gain of the mutation in the metastatic disease.

In CRB63 point a) was observed where a mutation was missed by the GOA panel. The designed GOA panel was relatively small, covering 148 hotspots, and the designed amplicons did not cover the entirety of each gene. Therefore, truncal mutations could be

missed if they occurred in regions or genes not covered by these amplicons, as was the case for the *CDK2NA* mutation in patient CRB63.

In both patients point b) was observed where tumour evolution resulted in a change in the driver mutation. In CRB55 a loss of a specific driver mutation was demonstrated and in CRB63 a gain of a mutation. Similar patterns of tumour evolution have been seen in other studies on GOA (Murugaesu et al., 2015). Therefore, due to tumour evolution, driver mutations could be missed by only sequencing the primary tumour, or by tracking a single mutation in the blood.

However, overall there is a low chance that the dominant / truncal mutation has been missed in many patients in this study. This is because analysis of all the other patients with detectable ctDNA relied on a single high frequency mutation which was tracked through the blood samples and provided clinical utility. On review of the data from cancerSEEK a *CDK2NA* mutation was seen in just 2 of 81 patients (2%) (Cohen et al., 2018) showing how low in frequency the missed mutations are likely to be. WES of the tumours, or use of targeted panels with a larger number of amplicons, covering more genes would help answer the question of how common undetected SNVs are in these tumours, and would likely be essential to improve clinical utility of ctDNA tracking. If such panels could also be used to sequence total cfDNA, they may also reveal how common an evolutionary change in the driver mutation is in GOA.

2. The cancer had no dominant mutation (SNV) and is driven by an SNCA or epigenetic mechanism

This is a possibility given that of the four patients treated with curative intent, with no detectable mutation in their tumour (CRB78, 94, 102, 109), three had a detectable SCNA following NanoString analysis (see section 5.7). GOA is also known to be a cancer that has high rates of SCNA and therefore it is likely that a number of the cancers, including those

with detected SNVs have SCNAs. Detection of ctDNA resulting from tumours with only SCNA is however more difficult than detection of SNVs.

3. The tumours in these patients did not produce ctDNA.

The overall levels of total cfDNA were relatively low for four of these patients (<2700 copies / ml) despite having recurrent metastatic disease, in keeping with the idea that they may simply not shed ctDNA. However, there are a number of patients in this study that have both low levels of total cfDNA and detectable ctDNA. Patients CRB74, and 121 had total cfDNA levels up to 7,000 copies/ml suggesting that points 1 and 2 are more likely to be responsible for the lack of ctDNA detection. However, ctDNA is not the only cause for raised total cfDNA, which can also be raised in inflammatory conditions or after exercise (Thierry et al., 2016). As it is difficult to prove the non-existence of ctDNA this remains a theory of exclusion.

4. Another difference links these patients

Patient CRB63 had their single site of metastatic disease resected and their primary oesophageal tumour treated with a radical dose of radiotherapy and thus may have no detectable ctDNA because they have no residual cancer. Another possibility is that localised disease does not release ctDNA as efficiently as metastatic disease. In keeping with this CRB71 and 121 had local relapse or relapse predominately to regional lymph nodes and this may be responsible for the lack of detectable ctDNA. CRB80 was also unusual in that relapse was to a muscle in the leg, and thus may also have had limited ability to shed ctDNA. This leaves only CRB74 that had “classic” metastatic disease amongst these ctDNA negative patients, and therefore it is this patient that points 1 or 2 are most likely to relate to.

5.9.2 Minimal residual disease

In total three patients had detectable post-surgical ctDNA; CRB55, 61 and 137. Two of these patients (CRB55 & 61) relapsed following their surgery, whereas the third (CRB137) remains

well 34 weeks (237 days) following surgery. Post-surgical ctDNA suggests the presence of minimal residual disease, as the detected ctDNA must be coming from undetected residual disease. However, it is not clear as to whether all patients with ctDNA postoperatively will relapse.

The disappearance of detectable *TP53* p. T125T in blood sample 2 of CRB137 is an important point. There are a number of reasons why this pattern may have occurred without intervention.

1. The blood sample was taken too early following surgery – 4 days post op, before the residual cancer ctDNA could disappear?

This is possible as the rest of the patients in this study had blood samples taken 8-12 weeks post surgery or later, this being the earliest time point at which further anticancer therapy could be undertaken. For CRB137 the blood was taken at 4 days, as the patient was 90 years old and not due to return to hospital after discharge. Therefore, the time period for residual ctDNA to disappear in the postoperative period was short compared to other patients in the study. Secondly with the highest VAF (88.8%) in the study the ctDNA may have been easier to detect at low levels. However, the half-life of ctDNA has been calculated to be 15 minutes to 2 hours (Lo et al., 1999, Khier and Lohan, 2018, Beiter et al., 2011), therefore it is unlikely that 4 days after surgery residual ctDNA from the cancer that was removed remains. The result is therefore likely a true result.

2. Has the minimal residual disease disappeared?

Proving the disappearance of a disease that cannot otherwise be detected is extremely difficult. Some non-viable tissue may have remained post surgery that has since been removed from the body, for example by the immune system, but there is little precedence for this in other cancers and none in GOA. Other patients in this study have had

undetectable levels of ctDNA and gone onto relapse or have progression of disease. Therefore, absence of ctDNA is not synonymous with absence of disease.

3. Has the disease become quiescent and may yet result in a relapse?

This is certainly a possibility, and further follow-up of this patient will help to clarify the answer to this question. If the patient remains relapse free at 5 years the patient may be considered cured, but the patient is 90 years old and may not survive that long.

5.9.3 Lead interval to relapse or progression – *A window of opportunity?*

In keeping with this idea of monitoring patients CRB62, 75 and 89 (section 5.6.1.1) had rising ctDNA levels before diagnosis of progressive disease. This occurred 53 days before relapse in the case of CRB89, and could have been longer in the other two patients (3 and 10 days) had blood tests been taken earlier. This suggests ctDNA levels can predict disease progression.

A similar lead time was seen before relapse in patients CRB55 and CRB61 (section 5.6.2.2), where evidence of minimal residual disease was detected within a lead time of 33 and 18 weeks (233 and 128 days) before relapse.

This raises an important clinical question as to how this knowledge helps to improve the care of patients with GOA i.e. whether intervention with anti-cancer therapy during this lead time presents a window of opportunity that may be beneficial to patients?

The standard of care for second line therapy in the UK is docetaxel for metastatic patients (Ford et al., 2014). If relapse is detected earlier docetaxel could also be started earlier. However, a clinical trial would need to be constructed to ensure that such an intervention was beneficial to patients and did not increase the amount of time spent on treatment to no survival advantage. Indeed, as patients were often well when these blood samples were

taken there is a risk that quality of life could be decreased by treating earlier due to a rising ctDNA level.

In the case of patients who are undergoing treatment with curative intent, the presence of ctDNA in the postoperative is a poor prognostic sign. The finding in CRB137 of a positive postoperative ctDNA sample which subsequently became negative without treatment, suggests that at least two consecutive blood samples would need to be confirmed positive before any additional intervention should be instigated. It is however important to arrange a clinical trial to prove whether a second line of chemotherapy has the ability to eradicate this minimal residual disease and improve survival.

Although eradication of ctDNA may be taken as a positive sign, many patients in this study have relapsed following ctDNA negative blood samples and thus we know that a negative blood sample will not be sufficient to ensure a cure. In leukaemia treatment there is the concept of a complete cytogenic response, where no cancer cells can be viewed on a blood film AND a complete molecular response where no evidence of genetic alterations specific to the cancer can be found via PCR analysis (Colombat et al., 2006, Terwilliger and Abdul-Hay, 2017). Even a complete molecular response is not a guarantee of a cure, and up to two years of additional therapy must be instigated to improve cure rates in childhood leukaemia (Nachman et al., 1998). Whilst I do not advocate such prolonged chemotherapy in patients with GOA who cannot tolerate such treatment intensity it should be considered that eradicating ctDNA should be seen as similar to a complete molecular response, i.e. it indicates disease reduction but is not synonymous with a cure. Treatment can and should continue past this point. The degree to which this should be done can only be answered through the creation of well a designed clinical trial.

Finally, the findings on prediction of relapse suggest that those patients who have a ctDNA positive pre-surgery blood sample are also at high risk of disease relapse in addition to those with postoperative positive blood samples. Therefore, such patients should also be included

in trials designed using ctDNA to define high risk groups of patients where further intervention may improve cure rates.

5.9.4 PEACE

The PEACE study has provided a valuable metastatic tumour tissue from two patients for further investigation of tumour genetics. WES of DNA from the metastatic samples and from diagnosis (OGD Biopsy) and curative surgery may help identify further mutations or SNCAs that may be traced in the total cfDNA samples. In the case of CRB58 it may help detect evidence of disease in total cfDNA earlier and more definitely than the KRAS p.*G12D* mutation was able to predict (Figure 57, page 199). I suspect ctDNA may be present in the pre-surgical sample of patient CRB58 (blood 1) due to high levels of total cfDNA (6212 copies/ml), which in this study were frequently associated with high levels of ctDNA. This assumes that the most prevalent mutation in the cancer was not the *KRAS* p.G12D mutation, and that WES can reveal the undetected mutations.

Analysis of samples from CRB69 may show how the cancer changed as it became less responsive to chemotherapy and showed more aggressive growth. In particular for CRB69 it would be interesting to find a difference between bloods 2 and 4 (both negative for *TP53* p.E271X) that may have suggested the more imminent return of the disease following blood 4 (Figure 47, page 180)

5.9.5 Brain disease

In this study two patients were confirmed to have metastatic brain disease, CRB116 and CRB69. Both patients had detectable ctDNA, but in patient CRB116 the ctDNA did not return at the time of brain relapse whilst in patient CRB69 the ctDNA did return. There is evidence that tumours within the brain may not release ctDNA efficiently into the circulation, most likely because of the effect of the blood brain barrier (Wang et al., 2015b, De Mattos-Arruda et al., 2015a, Bettegowda et al., 2014) and this is the most likely reason for ctDNA not being detectable in patients with brain tumours.

Patient CRB116 had detectable ctDNA at diagnosis when the primary cancer was in situ. However, at relapse patient CRB116 had no disease outside the brain, and this may explain why there was no detectable ctDNA directly before, or at any time after relapse. Patient CRB69 in contrast had dynamic ctDNA levels and presence of ctDNA shortly after relapse to the brain. At the point of brain disease relapse CRB69 also had lymph node progression, meaning the ctDNA detected may have been from the lymph node disease rather than brain disease. As CRB69 has also been recruited to the PEACE study, further analysis of his metastatic tumours and ctDNA may help decipher which tumours the ctDNA is predominantly derived from.

5.9.6 SCNA analysis and *ERBB2* amplification

Amplification of specific genes was demonstrated in 13 of 20 patients (65%) using NanoString nCounter of tumour DNA (Figure 70, page 222) and confirmed by ddPCR and qPCR in 3 patients. However, none of the three patients treated with curative intent had detectable SCNA in their total cfDNA samples (Figure 72, page 225).

In an expansion cohort of 14 HER positive patients *ERBB2* amplification was detected in tumour DNA from 8 patients (57.1%) (Figure 73, page 228). *GAPDH* was the most discriminative reference gene to detect *ERBB2* amplification in tumour DNA samples. The use of tumour DNA *ERBB2* CN to predict HER2 IHC positively yielded a sensitivity of 50.0% and a specificty of 87.5% with a pre-defined amplification cut-off of CN=4.2. This suggests that this method could determine HER2 status and warrants further investigation in a larger cohort. High concordance (78.5%) between tumour DNA and total cfDNA determination of *ERBB2* CN suggests that total cfDNA analysis could be used to determine HER2 status if tumour DNA is not available.

Clinically less than 50% of patients who are HER2 positive respond to trastuzumab treatment (Bang et al., 2010). It is therefore important to evaluate whether the use of *ERBB2* amplification in tumour DNA or total cfDNA may be superior to IHC/FISH in selecting patients who will respond to trastuzumab. There are two reason to believe that *ERBB2*

amplification may be a superior biomarker than IHC/FISH alone. Firstly, amplification of *ERBB2* is linked with over transcription of *ERBB2* mRNA and overexpression of HER2 protein and response to trastuzumab (Seo et al., 2014, Bang et al., 2010). Secondly, there is clinical precedence for the link between gene amplification and response to trastuzumab, since IHC HER2 2+ tumours are assessed via FISH analysis for gene amplification, and such FISH positive patients have been selected for HER2 targeted therapy in clinical trials (Bang et al., 2010). As a result, FISH determination of ERBB2 amplification is now used in routine clinical practice to determine sensitivity to trastuzumab in patients with HER2 2+ tumours.

Six of 14 (42.8%) HER2 positive patients had no detectable *ERBB2* amplification in their corresponding tumour (Figure 73, page 228). This suggests that the cause of their overexpression of HER2 is due to a cause other than *ERBB2* amplification and is most likely epigenetic in nature (Singla et al., 2017). The response of this group of HER2 IHC/FISH positive, *ERBB2* unamplified patients to trastuzumab is currently unclear, and would be of interest to link with clinical responses.

A single biopsy may not accurately reflect overall *ERBB2* amplification due to intratumoural and temporal genetic heterogeneity (Gerlinger et al., 2012, de Bruin et al., 2014). Therefore, a patient with GOA may have some *ERBB2* unamplified tumours (resulting in a negative biopsy) and some positive tumours that result in an amplified total cfDNA result, due to detection of cfDNA shed from these amplified tumours. It is therefore plausible that patients with total cfDNA *ERBB2* amplified results may benefit from trastuzumab therapy even when their tumour DNA analysis is not *ERBB2* amplified. Further investigation of this in clinical trials would be required to prove this theory.

5.9.7 Overview

This study has shown that ctDNA was readily detectable by SNV analysis in the majority of patients with GOA. Dynamic changes in ctDNA levels tracked treatment response and resistance to anti-cancer therapy, and high levels of ctDNA at relapse predicted poor survival. The detection of ctDNA in patients undergoing curative treatment was linked with poor relapse free survival, and when present in post-surgical blood samples it showed evidence for the presence of minimal residual disease, identifying patients at high risk of relapse.

The detection of SCNAs in total cfDNA was more challenging. Its presence was shown by *ERBB2* copy number analysis in metastatic HER2 positive patients. This could allow the use of total cfDNA analysis to determine which patients receive trastuzumab based therapies. However, the study cohort was small and must be expanded, before it can be used clinically. In contrast the use of SCNA to detect minimal residual disease does not appear to be possible using the methods described.

6 General Discussion: The translational use of circulating tumour DNA

This project has demonstrated that ctDNA can be detected in patients with both breast cancer and GOA. Whilst ctDNA has been detected previously in both types of cancer (Shoda et al., 2017, Hamakawa et al., 2014, Fribbens et al., 2016, Page et al., 2017, Shaw et al., 2012, Dawson et al., 2013), this study has shown an expanded role in both cancers. Tracking of tumour specific SNVs in serial total cfDNA samples was shown in both cancers, which reflected and predicted response of the disease to treatment. In breast cancer emergent *ESR1* mutations also predicted non response to oestrogen targeted therapies, suggesting that ctDNA monitoring could provide clinicians with valuable clinical information on the selection of appropriate therapies. These findings indicate the importance of ctDNA as an emerging biomarker

A similar pattern of ctDNA response has been seen in both cancers in metastatic disease, where during periods of response to therapy ctDNA became undetectable. However, this did not equate to disease eradication. In every metastatic patient ctDNA levels subsequently returned indicating a 'molecular relapse' that frequently preceded and therefore predicted progression of disease. In contrast when disease response was seen on CT scans and ctDNA levels did not become undetectable, response to anticancer therapy was short, especially in GOA. In both cancers the highest levels of ctDNA were seen at times of disease progression, particularly in the final sample before death. This suggests that if tumour profiling is desired from total cfDNA analysis, using samples taken at times of disease progression is most useful. This finding was demonstrated particularly well when WES of a total cfDNA from a breast cancer patient was undertaken, as the most informative bloods were taken at the points of most significant disease progression, whereas those taken at time of disease response yielded few somatic mutations.

In the neoadjuvant setting, the presence of detectable ctDNA in breast or GOA patients prior to surgery was identified as a poor prognostic sign, with all three of the GOA patients in which this finding occurred, having relapse of their disease. In contrast for breast cancer patients 2 of the five patients with this finding have so far relapsed, and this occurred in the two patients with the highest levels of ctDNA. It is probably that the difference between these two cancers in response to pre-surgical ctDNA is related to the difference in biology between these two cancers. GOA cancers tend to relapse early following surgery (Cunningham et al., 2006) and frequently most relapses have occurred within the two year period covered by this study. In contrast breast cancer relapses can occur 10 years (Fisher et al., 2002) or more after initial treatment with curative intent and therefore, those patients with pre-surgical ctDNA who have not yet relapsed are likely to be of high chance of relapse.

This study identified lag times in a number of patients between rises in ctDNA and clinical relapse or progression. In order to act on a rise in ctDNA levels it should first be necessary to define what a rise in ctDNA levels is. I therefore propose that a rise of 10 fold in ctDNA levels, or a change from undetectable to detectable ctDNA levels should be classified as a 'ctDNA rise' whereas all increases less than this to be defined as 'stable ctDNA' levels. The same would then therefore be true for a fall in ctDNA levels. To confirm that ctDNA levels are truly predictive of relapse far larger studies of are necessary with regular blood sampling of no less than 3 months in order to detect early relapses. Ideally a time period of 2 months or less would be necessary in early relapsing tumours such as GOA or triple negative breast cancer with longer periods between bloods for slow relapsing cancers such as ER+ breast cancers. If these studies could show that the majority of patients had rising ctDNA levels before relapse or progression of disease, this would provide strong evidence for prospective studies to be undertaken, with clinical interventions.

The study on GOA also demonstrated the importance of the post-surgical sample in determining relapse. In the three patients where this sample was positive for ctDNA two relapsed early, and one remains relapse free at 33 weeks (230 days). It is likely that the

presence of circulating tumour DNA in these samples reflects the presence of minimal residual disease as the mutation detected in the total cfDNA must be derived from residual cancer cells. However, it is clear that patients can have minimal residual disease and be ctDNA negative. Therefore, it may be more useful to consider positive ctDNA as a molecular relapse, that is highly likely to result in clinical relapse. In the future this could be used as a marker for intervention in clinical trials to determine if further treatment, taken upon the finding of molecular relapse, can alter the trajectory of disease.

The work on WES and cfDNA in one patient with breast cancer, is the first example of the use of WES to sequence multiple serial samples in a breast cancer patient. Analysis has provided significant additional insight into the genomic evolution of the patient's breast cancer, providing additional information into which therapies may be most effective and predicting resistance to HER2 targeted therapy. WES of cfDNA was most informative from samples taken at significant disease progression but even then was limited by depth of sequencing, missing rare yet clinical important mutations in *ESR1.* The main strength of WES is that it detects the majority of high frequency gene mutations and the main weakness lies with the large amount of data processing time and computer programming technology needed to decipher this large amount of information. This project offers initial insight into what can be learned from WES showing evidence of both scientific and clinical importance. As technology improves we will be able to increase the depth of WES and also under take WGS of total cfDNA, both of which will provide many orders of magnitude more information about single time points in a patient's treatment.

Genomics England's flagship '100,000 genomes project' includes using WGS to sequence 50,000 genomes from patients' with cancer including tumour and germline DNA (England, 2018). This will provide whole-genome information on these cancers and identification of tumour specific mutations and SCNAs. These variants could be tracked in total cfDNA, although this is not part of the project. However, the 100,000 genomes project will likely be just the first step in the pathway to genomic information becoming widely available in the NHS. With the availability of this information tumour tracking using tumour specific ctDNA

mutation analysis may also become a more feasible proposition, using data from tumour WGS. This requires the clinical / translational utility of ctDNA tracking to be demonstrated, showing the importance of continuing the work highlighted by the projects in this thesis.

Continually evolving sequencing technology highlights an important issue in the development of ctDNA. Even within the 3-year time span of this project sequencing technology has improved as shown by the increased sensitivity of Oncomine technology in Chapter 3. However, even with improving technology there is always a trade-off between read depth and the degree of coverage of the genome. In general, the higher the degree of coverage of the genome the lower the read depth, and therefore the lower the sensitivity is for detection of rare mutations, such as those present in cfDNA (Abel and Duncavage, 2013). This is shown in this project by the low depth (mean depth 143) following WES cfDNA analysis compared to up to a read depth of 10,000 for the GOA and AcquiRes amplicons. In additional depth is affected by the amount and quality of available input DNA, with larger amount of high-quality DNA yielding better sequencing results with higher depth. It is therefore difficult to decide on the correct size of panel for all situations. In general, small panels of 5-50 genes provide better depth and reduce the chance of missing low frequency mutations (e.g. the *ESR1* mutations in CRB24) of the genes included (Stasik et al., 2018). However, if the mutations that are being searched for cannot be known in advance then high coverage of the genome such as WES is preferable(Ahlborn et al., 2019). Therefore, I would recommend that when cfDNA levels are high and the potential mutations are unknown WES and even WGS of cfDNA is preferable. However, when input DNA levels are limited and mutations in certain genes or gene areas can be predicted smaller high depth panels of involving <50 genes are preferred, and when very high sensitivity confirmation of mutations is required ddPCR is still useful.

It should also be noted in the GOA study that tumour specific SNVs were detected in cfDNA in a significant number of patients whilst the detection of SNCAs was much more limited. Although less detailed work was carried out in the SCNA section it is apparent that SCNAs are more difficult to detect in ctDNA and this is reported in the literature (Molparia et al.,

2017, Li et al., 2017). However, the ability to detect SCNAs also depends on the copy number of the amplification in the initial tumours, for example a copy number of 100 is more readily detected in ctDNA than one of 5. The analysis of ctDNA in the plasma of patients with tumour DNA containing SCNAS will be less sensitive than for cancers with one or more high frequency SNV, as one to a few molecules containing an SNV can be detected by ddPCR / Oncomine sequencing (Hardenbol et al., 2003, Li et al., 2018a), whereas the same is not true of SCNAs. It is therefore apparent that SCNA tracking (as for SNVs) will be most clinically useful in the palliative setting where ctDNA levels are higher. As different cancer types show different levels of SNVs and SCNAs the ease with which ctDNA can be detected will vary from cancer type to cancer type and in the palliative and adjuvant / neo-adjuvant setting and it is therefore difficult to predict which cancers will be most amenable to ctDNA analysis. Despite these findings it is clear that metastatic patients have increased ctDNA levels on average which allows increased ease of ctDNA detection in these patients. In addition, since SNVs are easier to track using high sensitivity, their use in the curative intent setting, where detection of low levels of ctDNA is critical, will be preferred in future studies.

6.1 Further work

In the three years since this study was embarked upon the technology available for NGS sequencing has evolved. In both the breast cancer and GOA, mutation specific ddPCR was required to detect low frequency mutations. However individual DNA barcoding of DNA molecules prior to NGS sequencing has allowed a significant improvement of the sensitivity of NGS sequencing in the last few years (Vollbrecht et al., 2018, Gale et al., 2018, Kivioja et al., 2011). This was demonstrated towards the end of the project using Oncomine™ technology. It is thus no longer necessary to design mutation specific ddPCR assays for mutations included within the coverage of a panel as the sequencing technology can now achieve comparable levels of sensitivity (0.1-0.05%). This will greatly improve the speed at which analysis of samples can be achieved, although with increased cost.

This study is the first to demonstrate the utility of tumour specific mutation tracking in total cfDNA in GOA. However, an expansion cohort is required to confirm the findings of this study particularly in the case of patients treated with curative intent, where only 3 patients out of 23 were shown to have ctDNA positive post-surgical samples. By utilising molecular barcoding of individual DNA molecule technology such as Oncomine™, it will no longer be necessary to design the individual ddPCR assays and therefore a larger number of variants across patient samples can be analysed in a shorter time. The main limiting factor will be recruitment of patients and obtaining bloods at regular intervals pre and post surgery, as well as funding the increased cost of each test. It will also allow multiple mutations to be assessed in each cfDNA at the same time, allowing an improved assessment of the dynamics of the evolution of these cancers over time, compared to single mutation profiling by ddPCR.

The same individual molecular profiling can be used to improve the detection of mutations in breast cancer. In this study more curative intent patients were recruited than analysed, and therefore more can be learned from this cohort, particularly in the analysis of the post-surgical samples to see if they reveal the same information as was seen in patients with GOA.

SCNA analysis in GOA has yielded a small number of patients with definitive detection of *ERBB2* amplification that can be detected in total cfDNA. Full analysis of all 44 patients by Nanostring for tumour specific SCNA may therefore yield further total cfDNA samples that have detectable tumour specific SCNAs. These may provide further clinical information on the nature of these cancers, and confirmation of the finding found during SNV analysis. Funding for this area of work has been jointly applied for and awarded to an MSc student to complete this work. However, as SCNA analysis relies on the detection of a skewed ratio rather than the presence or absence of a specific mutation, analysis in the total cfDNA is inherently less sensitive to the presence of ctDNA.

The finding that a single patient with squamous cell oesophageal cancer has very high levels of ctDNA is also an important finding as it reflects a similar finding in squamous cancers versus adenocarcinomas of the lung in the TRACERx study (Abbosh et al., 2017, Jamal-Hanjani et al., 2017). This suggests that a similar study to the GOA study run in squamous cell oesophageal cancer may yield good results.

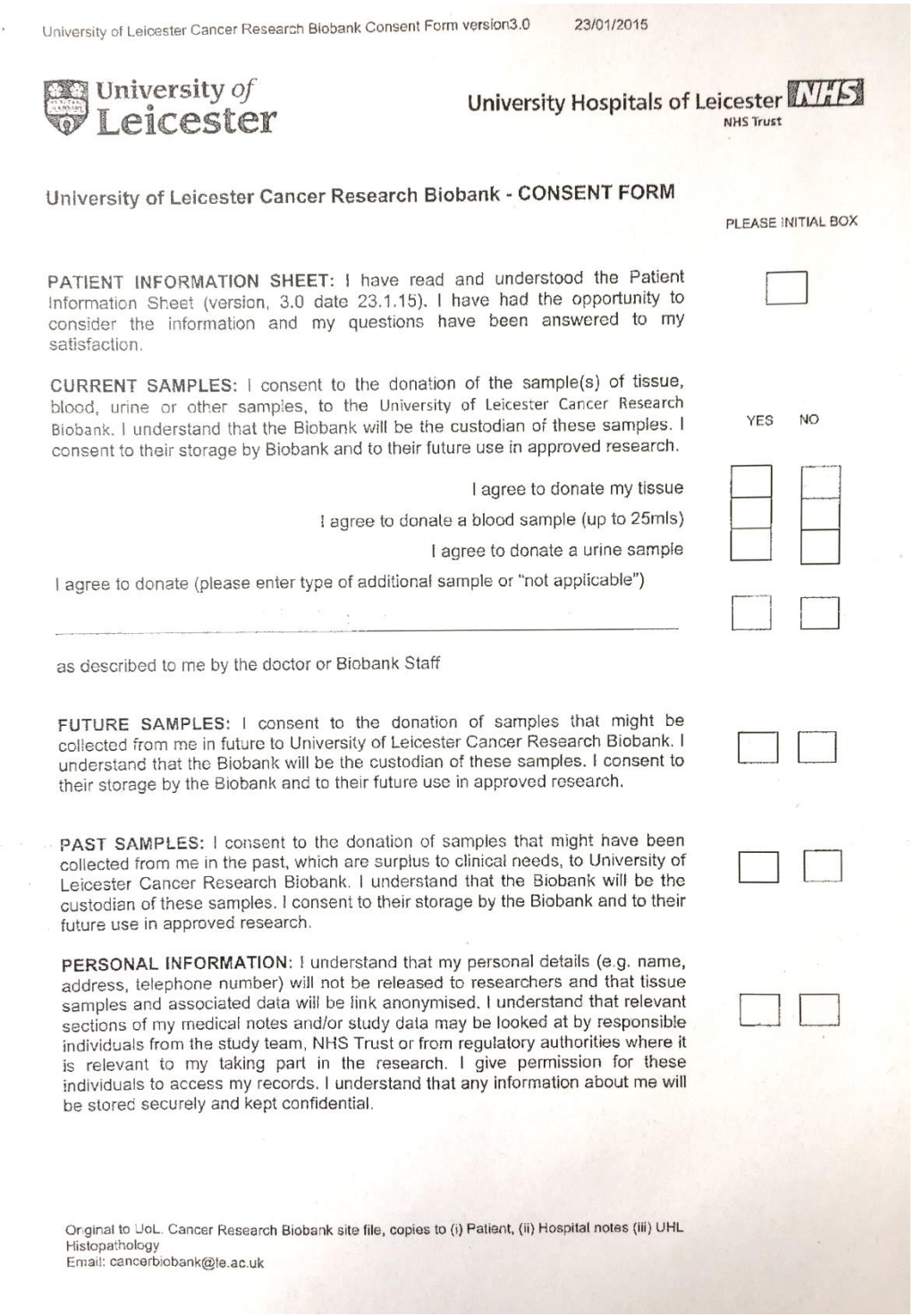
Samples from the PEACE study will also provide an important method for further investigating the tumour genetics of GOA. The aim is to test the hypothesis that WES analysis of sampled metastatic tumour DNA, will reveal the full extent of tumour specific mutations and indicate the degree of evolution that has occurred in these patients over time. Following this a targeted cfDNA NGS panel will be designed using the variants identified by WES, and the cfDNA samples already obtained reanalysed. This will provide an insight into the extent to which total cfDNA can be used in the future for personalised medicine. Funding has been awarded for this project.

A number of patients in this study affected by both breast cancer and GOA had metastases to brain. There is evidence that cfDNA does not pass the blood brain barrier (Bettegowda et al., 2014, De Mattos-Arruda et al., 2015b) but is present in the cerebrospinal fluid (CSF) of patients with metastatic brain disease (Momtaz et al., 2016) and primary brain tumours (De Mattos-Arruda et al., 2015b). It is therefore of interest to discover if CSF contains cfDNA that could better inform scientists and clinicians about the nature of brain tumours. Work by Dr Page in the group has shown evidence of polyclonal *ESR1* gene mutations in the CSF of a patient with MBC and brain mets (Karen Page, personal communication and A Hills PhD thesis, Imperial College 2018), however other confirmatory studies are needed. Therefore, if CSF could be obtained from patients with GOA or Breast cancer and brain tumours via PEACE, or from those patients with metastatic brain disease that routinely need CSF sampling and intracthecal chemotherapy such as lymphoma (Kwong et al., 2009), and choricarcinomas (Ngan et al., 2018) the use of ctDNA in CSF could be further investigated.

The detection of ctDNA has been shown to have a role as a predictive and prognostic blood based biomarker. Other biomarkers are already in widespread use such as βHCG and AFP in germ cell tumours, PSA in prostate cancer and Ca125 in ovarian cancer (Duffy, 2015). However, the true usefulness of even these biomarkers is not without controversy. For example, Ca125 was widely used to monitor ovarian cancer patients for relapse, but a well controlled study in 2010 (Rustin and Burg, 2009) showed that commencing palliative chemotherapy due to rising Ca125 levels did not improve overall survival and increased time spent on chemotherapy. Therefore, Ca125 is no longer used to make treatment decisions in isolation. Thus, in order for ctDNA to become widely used, well designed clinical trials are required where treatment decisions are based upon findings from ctDNA, to confirm that ctDNA is an accurate biomarker and that acting on ctDNA levels by commencing additional treatments improves survival and alters outcomes, rather than simply increasing time on treatment. A number of prospective trials are in progress to answer these questions such as plasmaMATCH (ICR, 2018) and c-TRAK TN (CRUK, 2018a) in breast cancer. Before similar trials can be embarked upon in a wider range of cancers, it must be demonstrated that these treatments decisions can be based on solid scientific evidence. This will require large cohorts of patients to donate blood samples for total cfDNA analysis during longitudinal studies and ongoing clinical trials, so that the full utility of ctDNA can be defined.

7 Appendix

7.1 Consent Form



University of Leicester Cancer Research Biobank Consent Form version3.0 23/01/2015

University of Leicester

University Hospitals of Leicester NHS
NHS Trust

University of Leicester Cancer Research Biobank - CONSENT FORM

PLEASE INITIAL BOX

PATIENT INFORMATION SHEET: I have read and understood the Patient Information Sheet (version, 3.0 date 23.1.15). I have had the opportunity to consider the information and my questions have been answered to my satisfaction. ☐

CURRENT SAMPLES: I consent to the donation of the sample(s) of tissue, blood, urine or other samples, to the University of Leicester Cancer Research Biobank. I understand that the Biobank will be the custodian of these samples. I consent to their storage by Biobank and to their future use in approved research.

	YES	NO
I agree to donate my tissue	☐	☐
I agree to donate a blood sample (up to 25mls)	☐	☐
I agree to donate a urine sample	☐	☐
I agree to donate (please enter type of additional sample or "not applicable") _____	☐	☐

as described to me by the doctor or Biobank Staff

FUTURE SAMPLES: I consent to the donation of samples that might be collected from me in future to University of Leicester Cancer Research Biobank. I understand that the Biobank will be the custodian of these samples. I consent to their storage by the Biobank and to their future use in approved research. ☐ ☐

PAST SAMPLES: I consent to the donation of samples that might have been collected from me in the past, which are surplus to clinical needs, to University of Leicester Cancer Research Biobank. I understand that the Biobank will be the custodian of these samples. I consent to their storage by the Biobank and to their future use in approved research. ☐ ☐

PERSONAL INFORMATION: I understand that my personal details (e.g. name, address, telephone number) will not be released to researchers and that tissue samples and associated data will be link anonymised. I understand that relevant sections of my medical notes and/or study data may be looked at by responsible individuals from the study team, NHS Trust or from regulatory authorities where it is relevant to my taking part in the research. I give permission for these individuals to access my records. I understand that any information about me will be stored securely and kept confidential. ☐ ☐

Original to UoL. Cancer Research Biobank site file, copies to (i) Patient, (ii) Hospital notes (iii) UHL Histopathology
Email: cancerbiobank@le.ac.uk

GENETIC ANALYSES: I consent to genetic analysis being done on my samples, including whole genome sequencing, to determine whether genetic makeup has any influence on disease. I understand that I will not receive any information on any genetic analyses unless they are pertinent to my care. ☐ ☐

ANIMAL RESEARCH: I consent to my samples being used in projects that use animals. ☐ ☐

COMMERCIAL COMPANIES: I understand that some these projects may be carried out under appropriate contract by researchers working for or with commercial organisations including the pharmaceutical industry and I agree to this. I understand that I will not benefit financially if research on my donated samples and data leads to new treatments or medical tests. ☐ ☐

FREEDOM TO WITHDRAW: I understand that I am free at any time to withdraw my consent for University of Leicester Cancer Research Biobank to store and use my samples without giving any reason and without it affecting my medical care. I understand that my samples would then be destroyed and my data will no longer be used. I also understand that any data from research already performed cannot be withdrawn. ☐ ☐

Person taking consent (Print): _____ Signature: _____ Date:_____

I have explained to the donor the reasons for collecting, storing and using samples for research. I am satisfied the donor signing this form understands the content and purpose of this consent form.

Donor name (Print): _____ Signature: _____ Date:_____

Original to UoL. Cancer Research Biobank site file, copies to (i) Patient, (ii) Hospital notes (iii) UHL Histopathology
Email: cancerbiobank@le.ac.uk

7.2 qPCR Primers

Primer	Forward	Reverse	Length
ERBB2	ATAACACCCACCTCTGCTTCGT	GGTGCGGGTTCCGAAAG	63
RPPH1	CGGAGGGAAGCTCATCAGTG	GACATGGGAGTGGAGTGACA	62
GAPDH	GGCTAGCTGGCCCGATTT	GGACACAAGAGGACCTCCATAAA	95
CNTNAP1	ACCTGGATGCGCTATAACCTACA	GGCTCAGCATGTGGGAGAAC	70

Table 46 – **DNA sequences and length of amplicon for primers used in ddPCR**

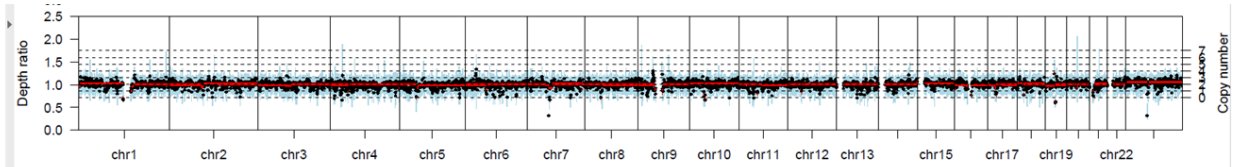
7.3 CRB77 *TP53* p.R175H mutation analysis

Timepoint	ddPCR Droplets Positive	Negative	VAF	NGS Coverage	Mutant call	VAF
1	89	944	8.3	1120	56	5.2
2	1	1343	0.07	7491	4	0.05
3	1	1818	0.05	1192	5	0.41
4	3	1071	0.27	Not sequenced		
5	19	1278	1.4	Not sequenced		
6	31	1789	17.6	Not sequenced		

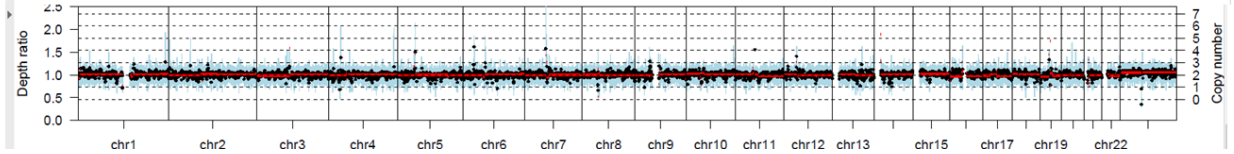
Table 47 – **Summary of ddPCR and NGS analysis of TP53 R175H mutation analysis of CRB77 plasma samples**

7.4 WES SCNA analysis in samples with low mutational burden

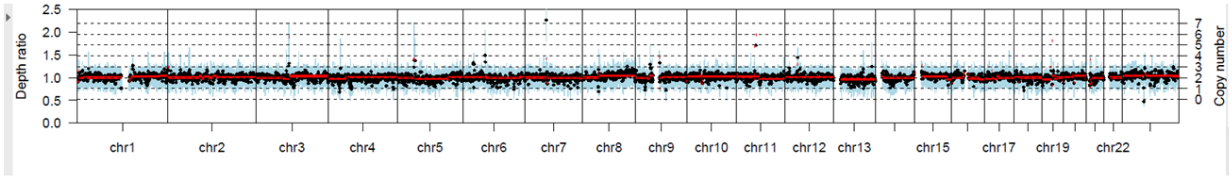
B14-15



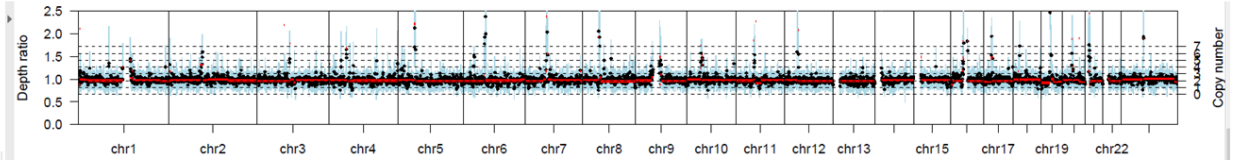
B31-15



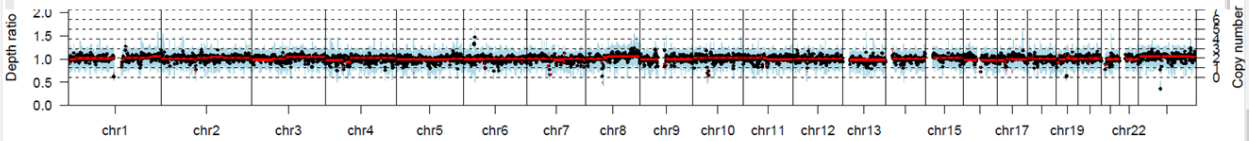
B14-16



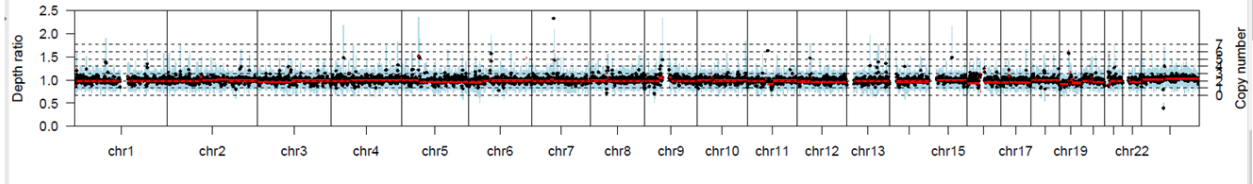
B44-16



B91-16



B143-16



7.5 WES – Sample cellularity, ploidy level, and BAF

Sample	Timepoint	Cellularity	Ploidy	BAF
T15-1	Initial	0.64	3.5	0.17
T16-1	relapse 1	0.28	2.1	0.16
T16-2	relapse 2	0.33	1.8	0.16
B14-15	1	0.299	1.9	0.17
B31-15	2	0.54	2	0.17
B90-15	3	0.76	1.8	0.16
B14-16	4	0.48	2	0.17
B44-16	5	0.31	2.2	0.17
B91-16	6	0.41	1.9	0.17
B143-16	7	0.32	2.1	0.17
B20-17	9	0.42	2.8	0.17

Table 48 **– Cellularity, Ploidy and B-allele frequency (BAF) for WES samples.**

7.6 Overview of ddPCR results

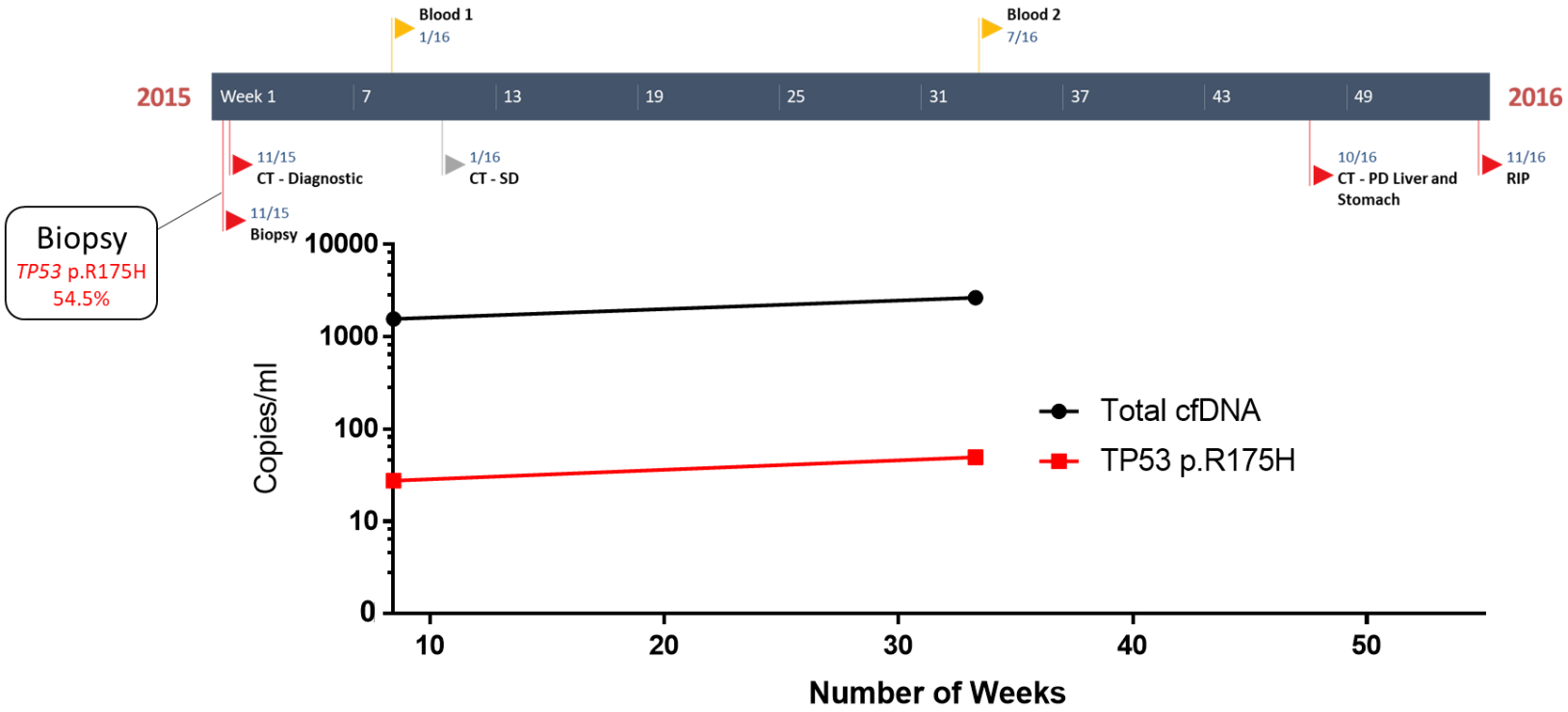
Curative Intent

		Blood 1					Blood 2					Blood 3					Blood 4					Blood 5					Blood 6				
Patient ID	Mutation	Sample	Mut	WT	Total	VAF	Sample	Mut	WT	Total	VAF	Sample	Mut	WT	Total	VAF	Sample	Mut	WT	Total	VAF	Sample	Mut	WT	Total	VAF	Sample	Mut	WT	Total	VAF
CRB055	R196X	B110/15	7	1077	11351	0.61	B016/16	1	2363	14562	0	B77/16	138	628	12253	17.8	B150/16	1643	2127	11658	43.1										
CRB058	KRAS G12D	B118/15	0	3138	12432	0	B031/16	0	2923	11585	0	B048/16	0	2144	11533	0	B142/16	0	1453	12247	0	B53/17	2	2812	12477	0.06	B60/17	19	2313	10648	0.72
		B93/17bld7	74	2505	10711	2.5	B113/18 Bld8	8	3966	13435	0.17																				
CRB059	C277F	B119/15	0	1586	15500	0	B30/16	0	2411	15478	0	B70/16		FAIL			B14/17	0	1906	15306	0	B54/17	0	2233	15894	0					
CRB061	R175H	B123/15	3	1714	11492	0.16	B037/16	0	1502	15270	0	B080/16	7	3373	15138	0.18															
CRB070	R213X	B15/16	0	2276	14194	0	B92/16	0	2362	13292	0	B47/17	0	2030	12299	0															
CRB071	R248Q	B19/16	0	624	11550	0	B11/17	1	1840	11541	0.05	B59/17	1	2616	11018	0.04	B62/18	1	1125	11510	0.08										
CRB080	R175H	B32/16	0	4479	13574	0	B60/16	0	4120	12937	0	B18/17	1	2760	12058	0.03	B46/17	0	2938	11173	0										
CRB082	R306X	B34/16	0	21763	11664	0	B78/16	0	2408	11472	0	B17/17	0	1467	11495	0															
CRB088	E286K	B51/16	0	1459	10629	0	B85/16	0	878	8245	0	B01/17	0	1398	1234	0	B48/17	0	1241	12690	0										
CRB095	P151R	B65/16	0	11918	112031	0	B131/16	0	1343	12908	0	B101/17	0	1260	13856	0															
CRB105	C135F	B089/16	6	1129	11790	0.5	B135/16	0	1751	11478	0	B10/17	0	1199	9356	0															
CRB107	Q192X	B097/16		SEQUENCED			B157/16		SEQUENCED																						
CRB110	C176F	B104/16	0	1970	11568	0	B149/16	0	1993	13123	0	B052/17	0	1808	12001	0															
CRB115	H179R	B124/16	0	1470	10689	0	B07/17	0	1990	12407	0	B061/17	0	2320	12591	0															
CRB116	G266V	B138/16	6	1199	14131	0.48	B04/17	1	1532	13277	0.06	B037/17	0	1677	10944	0	B79/17	0	1831	13305	0	B56/18	0	2164	11021	0					
CRB117	R213X	B139/16	0	1017	12084	0	B23/17	0	1916	13173	0	B75/17	1	2426	12718	0															
CRB137	T125T	B92/17	7	3415	12623	0.18	B206/18	0	376	9878	0																				
CRB144	S215R	B124/17	0	2211	15019	0																									
CRB145	E286K	B106/17	0	1685	11891	0																									
Palliative																															
CRB62	R280K	B124/15	30	1470	12110	1.9	B39/16	0	4129	14540	0	B100/16	83	3236	12960	2.19	B148/16	4	4252	14790	0.08	B06/17	805	6482	13514	8.6					
CRB63	R282W	B125/15	0	2486	11825	0	B42/16	0	1904	12360	0	B146/16	0	1655	11593	0	B012/17	2	5215	12103	0.029	B041/17	1	1297	12274	0.07					
CRB64	R248Q	B128/15	2	4138	14275	0																									
CRB67	R175H	B05/16	73	3449	12591	1.78	B79/16	86	3755	12107	1.88																				
CRB69	E271X	B09/16		Sequenced			B93/16	0	653	14506	0	B90/17	984	1217	13233	44.5	B54/18	1	3218	14094	0	B163/18	5	824	12041	0.6					
CRB74	S241A	B24/16	0	1322	12248	0	B61/16	0	856	9832	0	B102/16	0	1305	13272	0	B136/16	0	1480	14263	0	B156/16	0	1835	13357	0	B44/17	1	1577	13308	0
CRB75	R213X	B25/16	5	2426	12753	0.19	B54/16	0	1845	120152	0	B112/16	1	1307	11547	0	B141/16	0	1337	11851	0	B08/17	0	1196	11834	0	B62/17	49	1069	10881	4.2
CRB83	R249G	B38/16		Sequenced																											
CRB84	P151S	B40/16	18	4135	17284	0.38																									
	SMAD4																														
CRB85	R361S	B41/16		Sequenced																											
CRB89	F113C	B53/16	177	605	11481	22.3	B111/26	119	1547	11301	6.7	B130/16	653	1368	11238	31.6															
CRB118	R213X	B36/17	24	453	11089	4.9																									
CRB121	V122Dfs26X	B50/17	0	1583	13564	0	B47/18	0	1759	11791	0																				
Exceptional Responder																															
CRB81	R306X	B33/16	1	1711	11186	0.05	B121/16	0	1726	10649	0	B158/16	0	3026	11117	0	B45/17	0	2682	10889	0										
Squamous cell cancer patient																															
CRB56	R282W	B111/15	2573	757	111133	79.5	B061/16	2702	362	13730	89.5																				

Table 49 – **Complete GOA ddPCR results.** Mut = mutant droplets, WT = Wild-type droplets, Total = total droplets

7.7 Timelines for metastatic GOA patients

7.7.1 Patient CRB67



CRB67

1	2	Blood number	
1553	2636	cfDNA copies /ml	
1.8	1.9	AF (%)	*TP53* p.R175H
28	49	Mutant copies/ml	

Figure 77 - **Timeline for patient CRB67.** Showing clinical events in relation to total cfDNA copies/ml (black) and ctDNA as defined by detection of *TP53* p.R175H also shown as copies/ml (red) by ddPCR . SD = Stable disease, PD = progressive disease,.

7.7.2 Patient CRB83

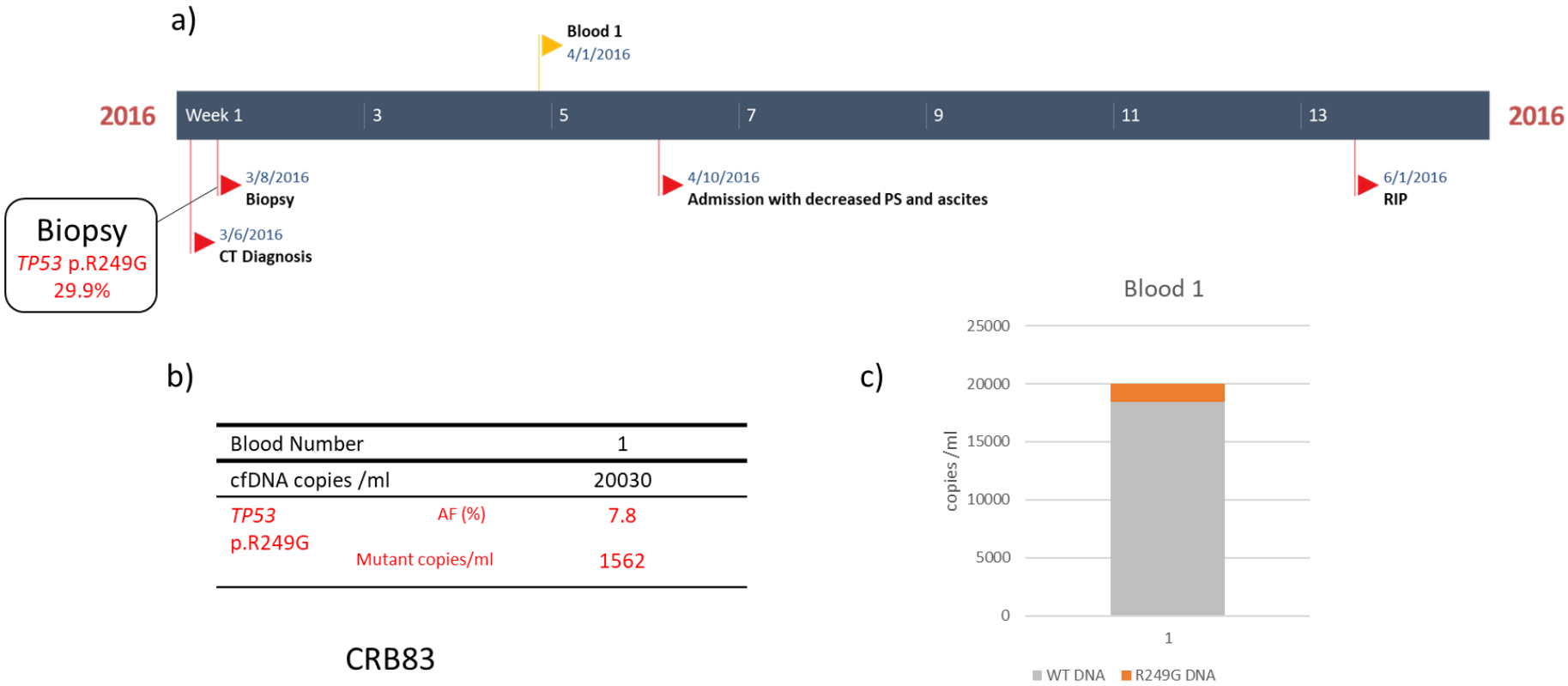
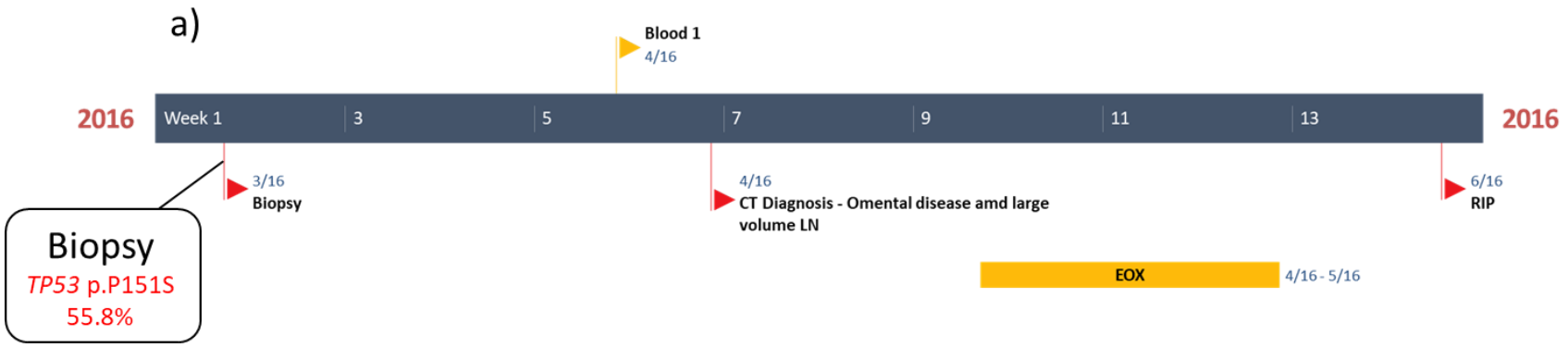


Figure 78 - **Timeline for patient CRB83** a) Showing clinical events. b) Summary table of total cfDNA ampliseq GOA sequencing results in copies/ml for the only blood sample (B38/16) c) Bar chart showing the ratio of the *TP53* p.R249G mutation to WT DNA at the same locus.

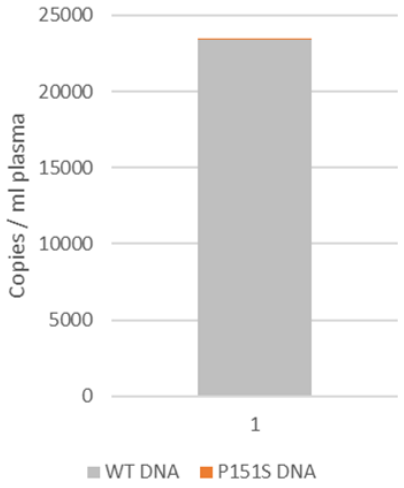
7.7.3 Patient CRB84



b)

Blood Number		1
cfDNA copies /ml		23484
TP53 p.P151S	AF (%)	0.38
	Mutant copies/ml	89

c)



CRB84

Figure 79 – **Timeline of CRB84.** a) Showing clinical events. b) Summary table of total cfDNA copies/ml and ctDNA by ddPCR in copies/ml for the only blood sample c) Bar chart showing the ratio of the *TP53* p.P151S mutation to WT DNA at the same locus.

7.7.4 Patient CRB85

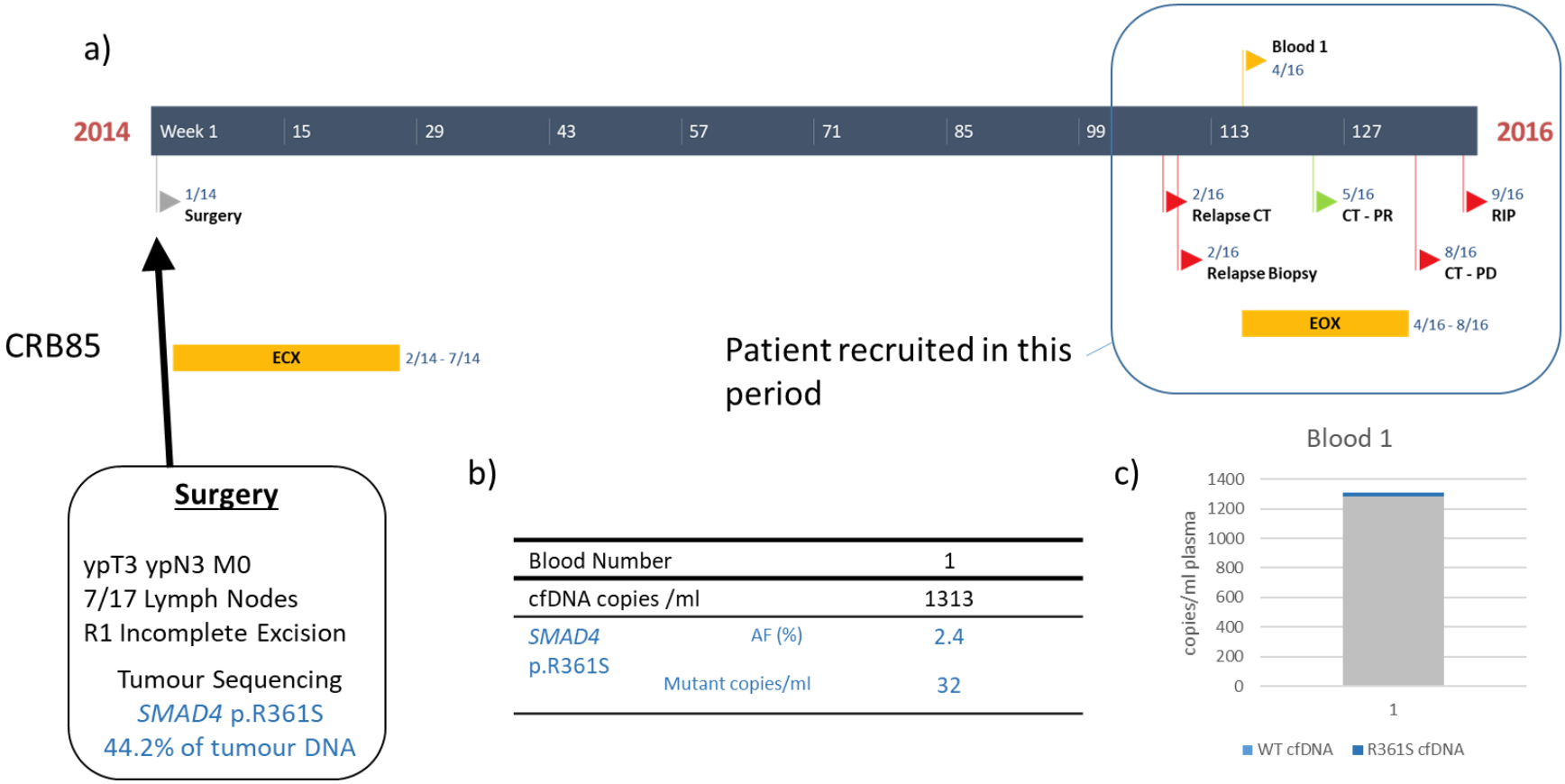
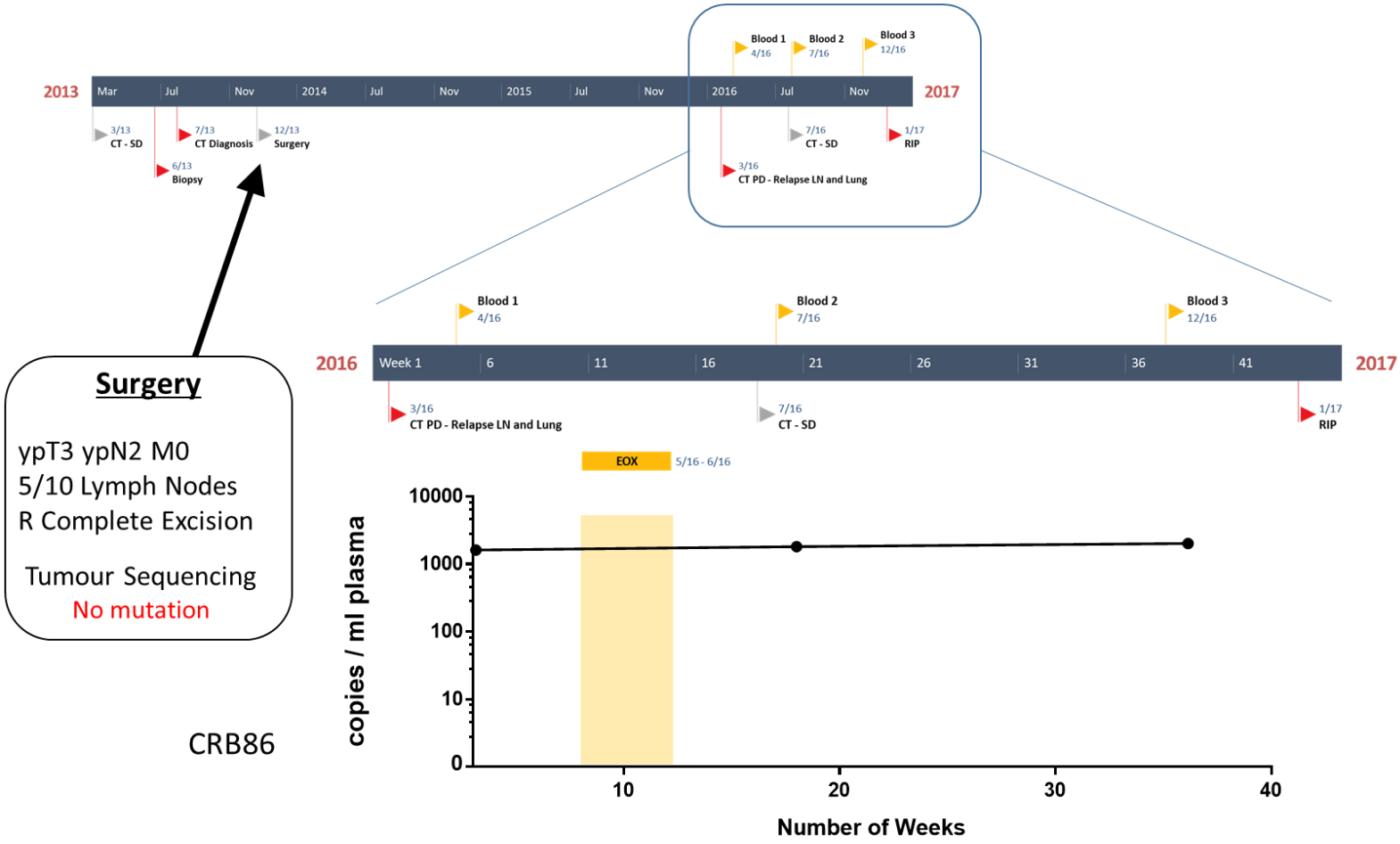


Figure 80 - **Timeline for patient CRB85** a) Showing clinical events and surgery preceding relapse in February 2016, when the patient was recruited into the study. b) Summary table of total cfDNA ampliseq GOA sequencing results in copies/ml for the only blood sample (B41/16) c) Bar chart showing the ratio of *SMAD4* p.R316S mutation to Wildtype (WT) DNA at the same locus. Sequencing details with a depth of 6193 at the p.R316S locus is shown in Table 37, page 174). PD = progressive disease, PR = Partial response.

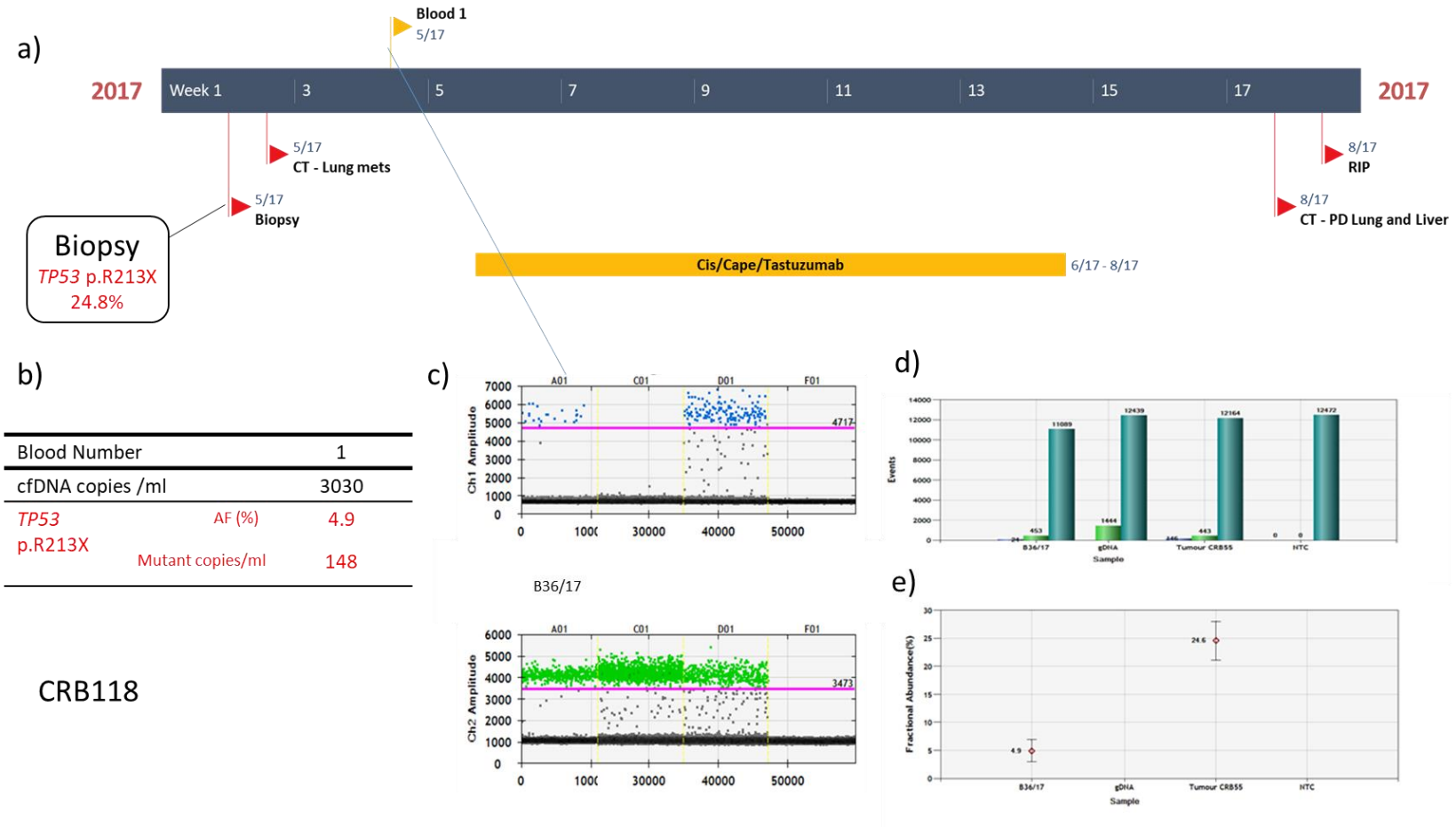
7.7.5 Patient CRB86



Blood Number	1	2	3
cfDNA copies /ml	1616	1818	2020

Figure 81 - **Timeline for patient CRB86.** Showing clinical events in relation to total cfDNA copies/ml (black). Top timeline is all clinical history. Main timeline from 2016 to 2017 represents the time following inclusion into study.

7.7.6 Patient CRB118



Blood Number		1
cfDNA copies /ml		3030
TP53 p.R213X	AF (%)	4.9
	Mutant copies/ml	148

Figure 82 - Timeline for patient CRB118. a) Showing clinical events, PD = progressive disease b) Summary table of total cfDNA copies/ml c) ddPCR results for *TP53* p.R213X for the only blood sample in lane A01, B36/17 showing positive droplets. d) The total number of events, and positive droplets for mutant and wildtype DNA. e) fractional abundance gDNA = genomic DNA, Tumour CRB55 = positive control DNA, NTC = non-template control.

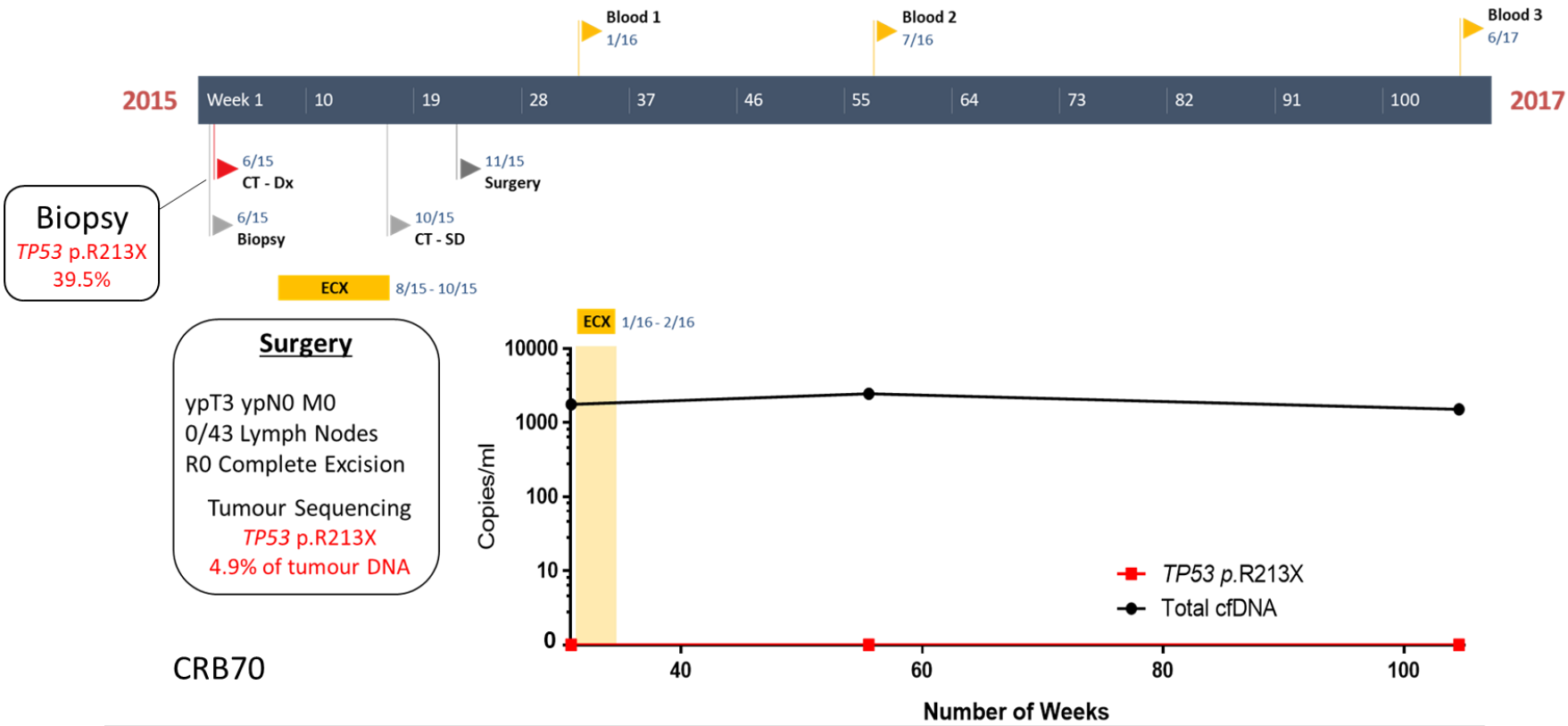
7.8 Timelines for GOA patients who did not relapse during follow-up and had trackable SNV

7.8.1 Patient CRB59

Blood number		1	2	3	4	5
cfDNA copies /ml		878	2583	1507	1674	2828
KRAS p.C277F	AF (%)	0	0	0	0	0
	Mutant copies/ml	0	0	0	0	0

Figure 83 - **Timeline for patient CRB59.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.C277F mutation not detected in the total cfDNA by ddPCR (red). SD = Stable disease.

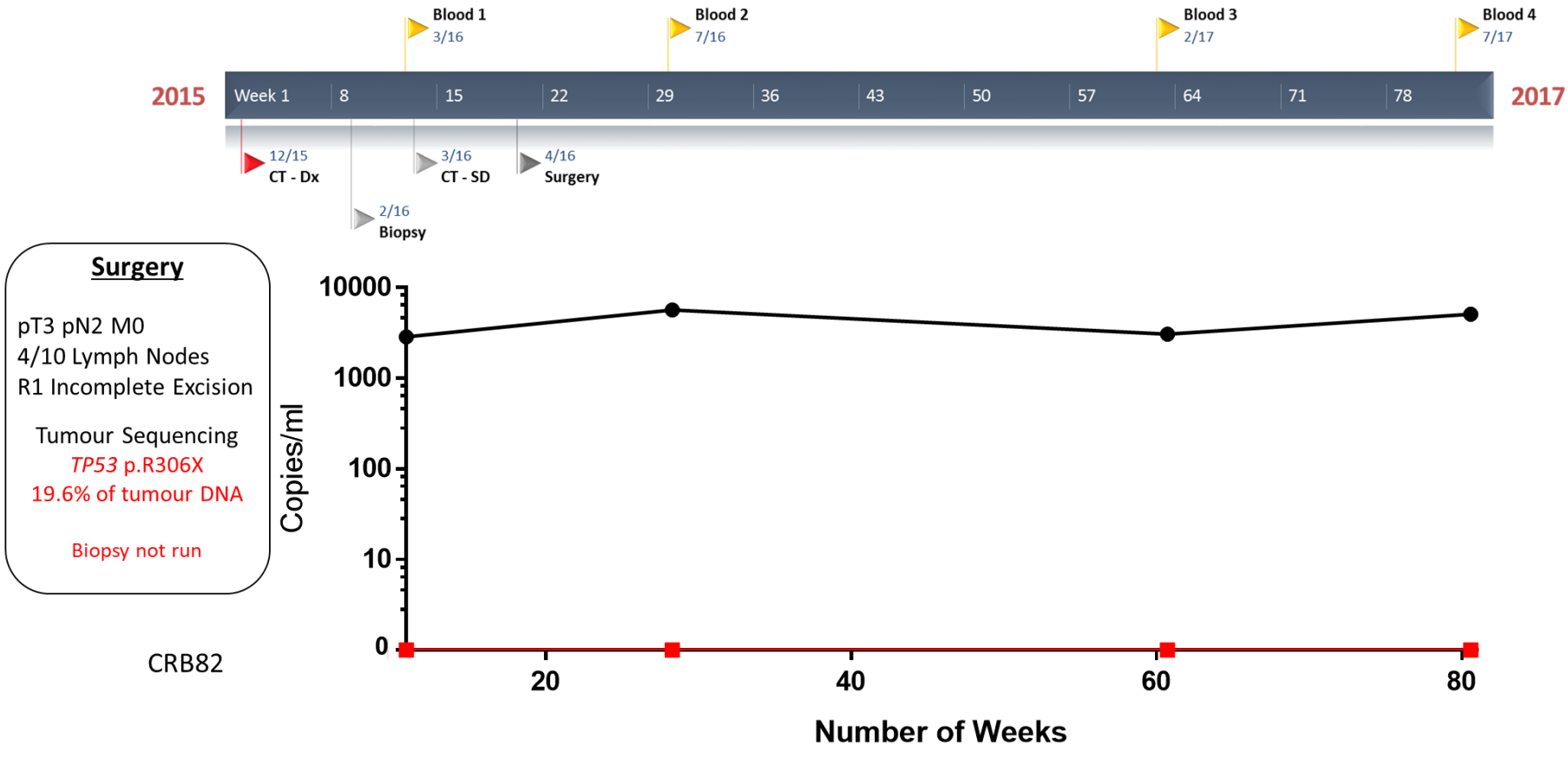
7.8.2 Patient CRB70



Blood Number		1	2	3
cfDNA copies /ml		1769	2462	1515
TP53 p.R213X	AF (%)	0	0	0
	Mutant copies/ml	0	0	0

Figure 84 - **Timeline for patient CRB70.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.R213X mutation not detected in the total cfDNA by ddPCR (red). SD = Stable disease.

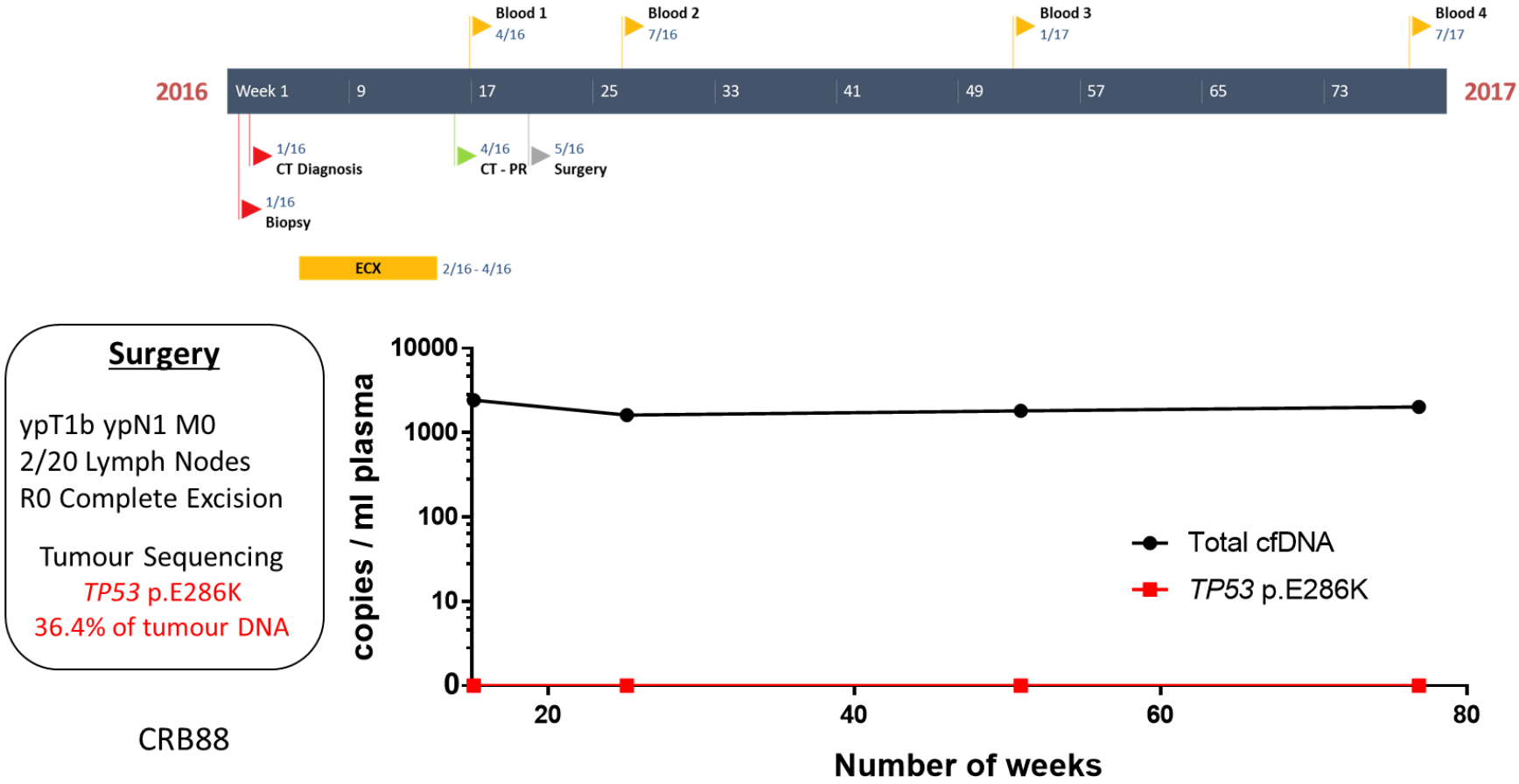
7.8.3 Patient CRB82



Blood Number		1	2	3	4
cfDNA copies /ml		2845	5627	3045	5066
TP53 p.R306X	AF (%)	0	0	0	0
	Mutant copies/ml	0	0	0	0

Figure 85 - **Timeline for patient CRB82.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.R306X mutation not detected in the total cfDNA by ddPCR (red). Dx = Diagnosis, SD = Stable disease.

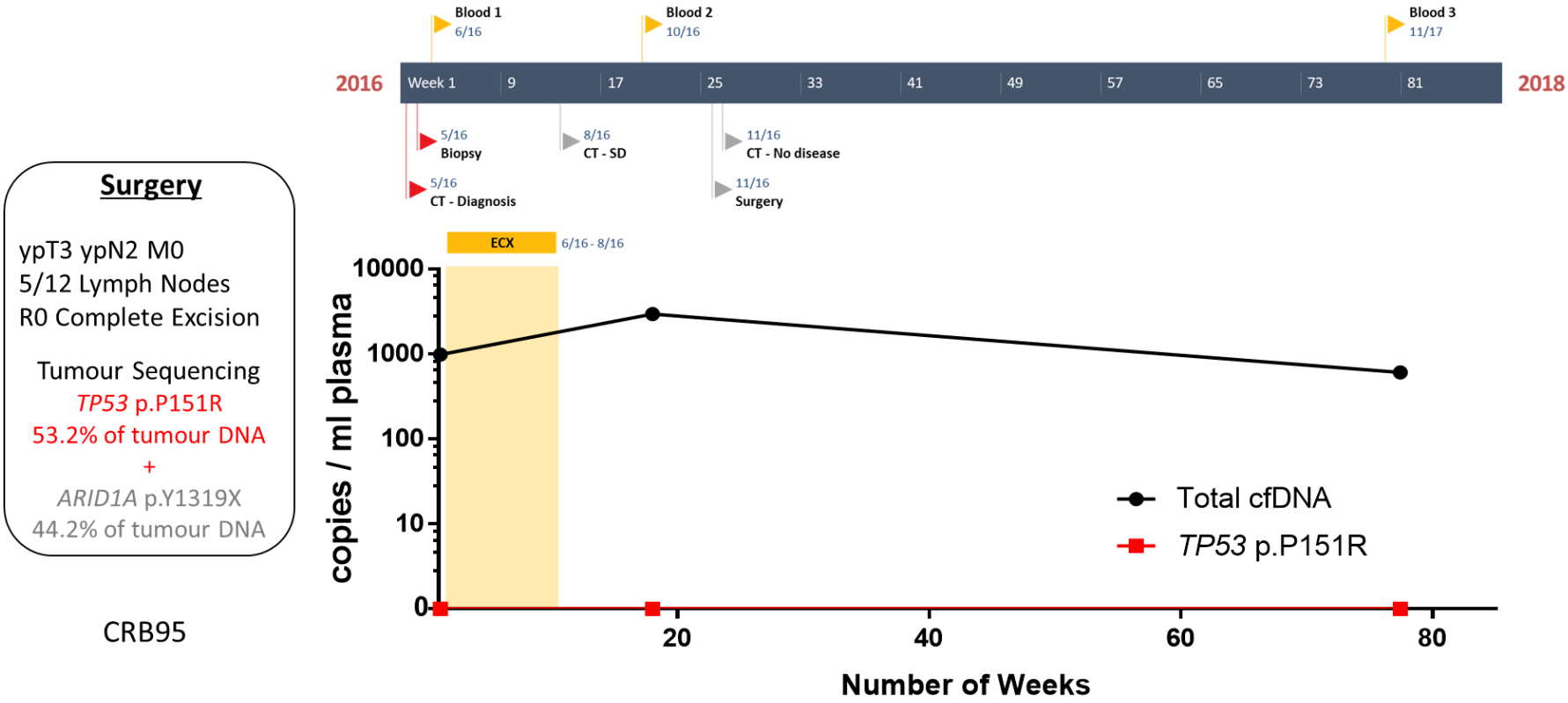
7.8.4 Patient CRB88



Blood Number		1	2	3	4
cfDNA copies /ml		2424	1616	1818	2020
TP53 p.E286K	AF (%)	0	0	0	0
	Mutant copies/ml	0	0	0	0

Figure 86 - **Timeline for patient CRB88.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.E286K mutation not detected in the total cfDNA by ddPCR (red)., PR = Partial response

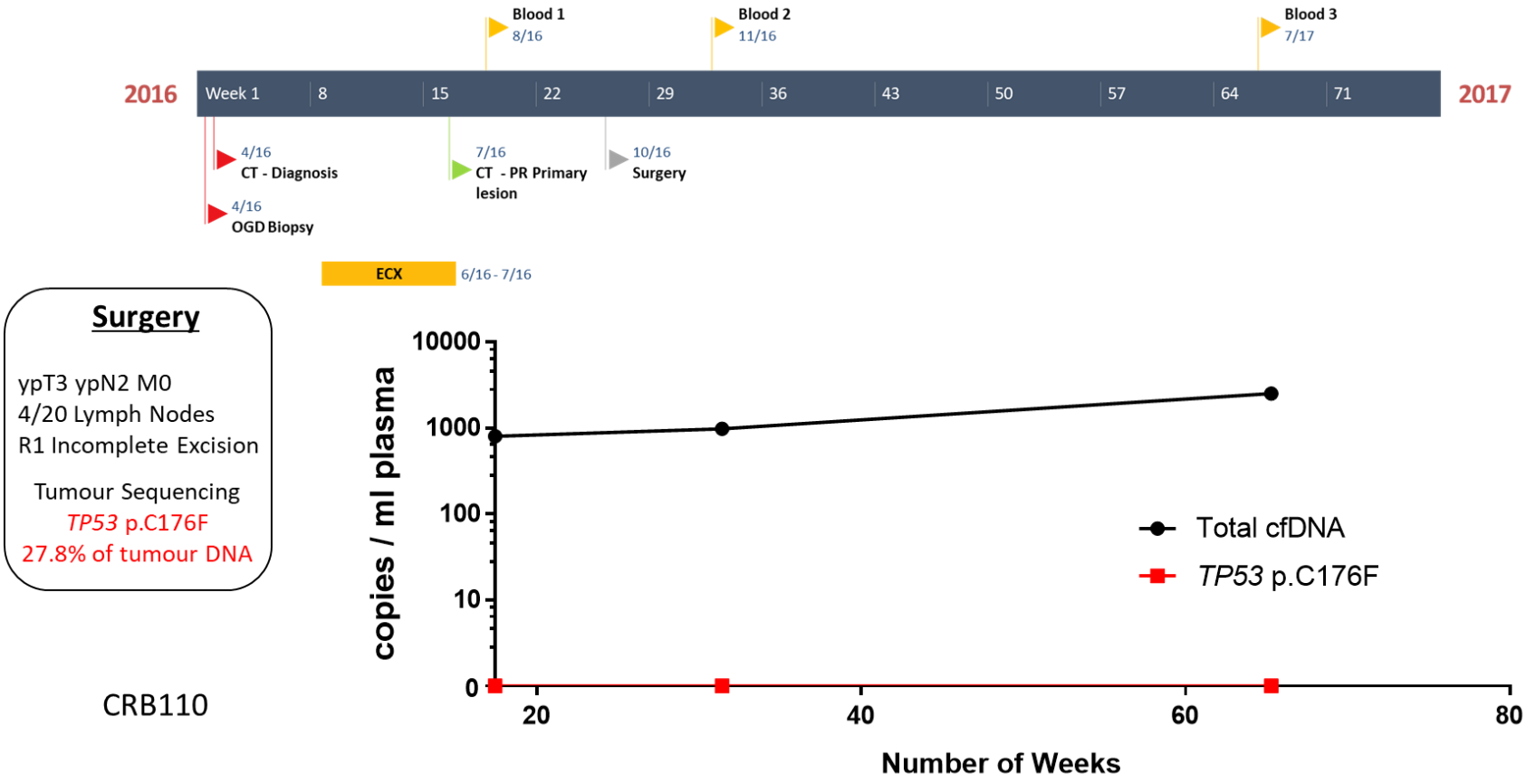
7.8.5 Patient CRB95



Blood number		1	2	3
cfDNA copies /ml		979	2929	606
TP53 p. P151R	AF (%)	0	0	0
	Mutant copies/ml	0	0	0
ARID1A p. Y1319X	AF (%)	0	Not done	Not done
	Mutant copies/ml	0		

Figure 87 - **Timeline for patient CRB95.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.P151R mutation not detected in the total cfDNA by ddPCR (red). The ARID1A mutation was not analysed by ddPCR as no assay was available. SD = Stable disease.

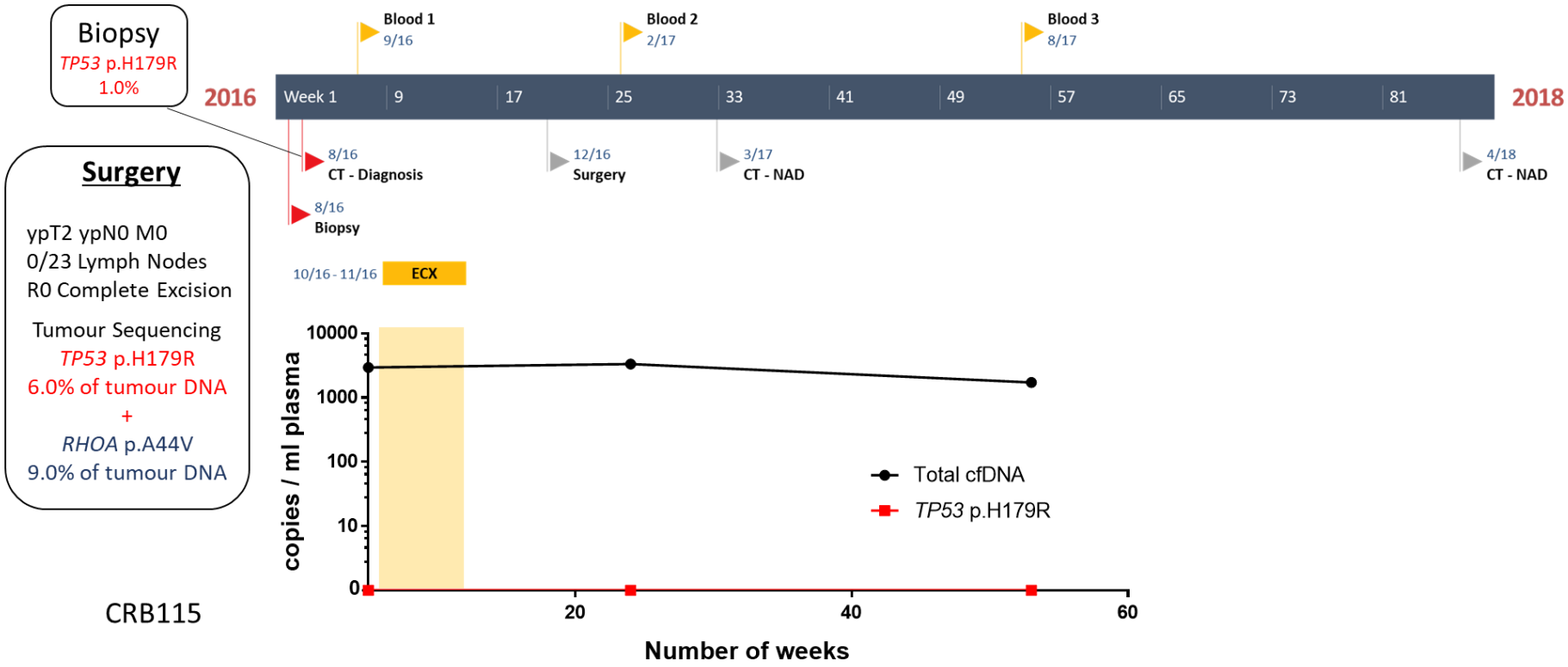
7.8.6 Patient CRB110



Blood number		1	2	3
cfDNA copies /ml		2930	3333	1717
TP53 p.C176F	AF (%)	0	0	0
	Mutant copies/ml	0	0	0

Figure 88 - **Timeline for patient CRB110.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.C176F mutation not detected in the total cfDNA by ddPCR (red). PR = partial response

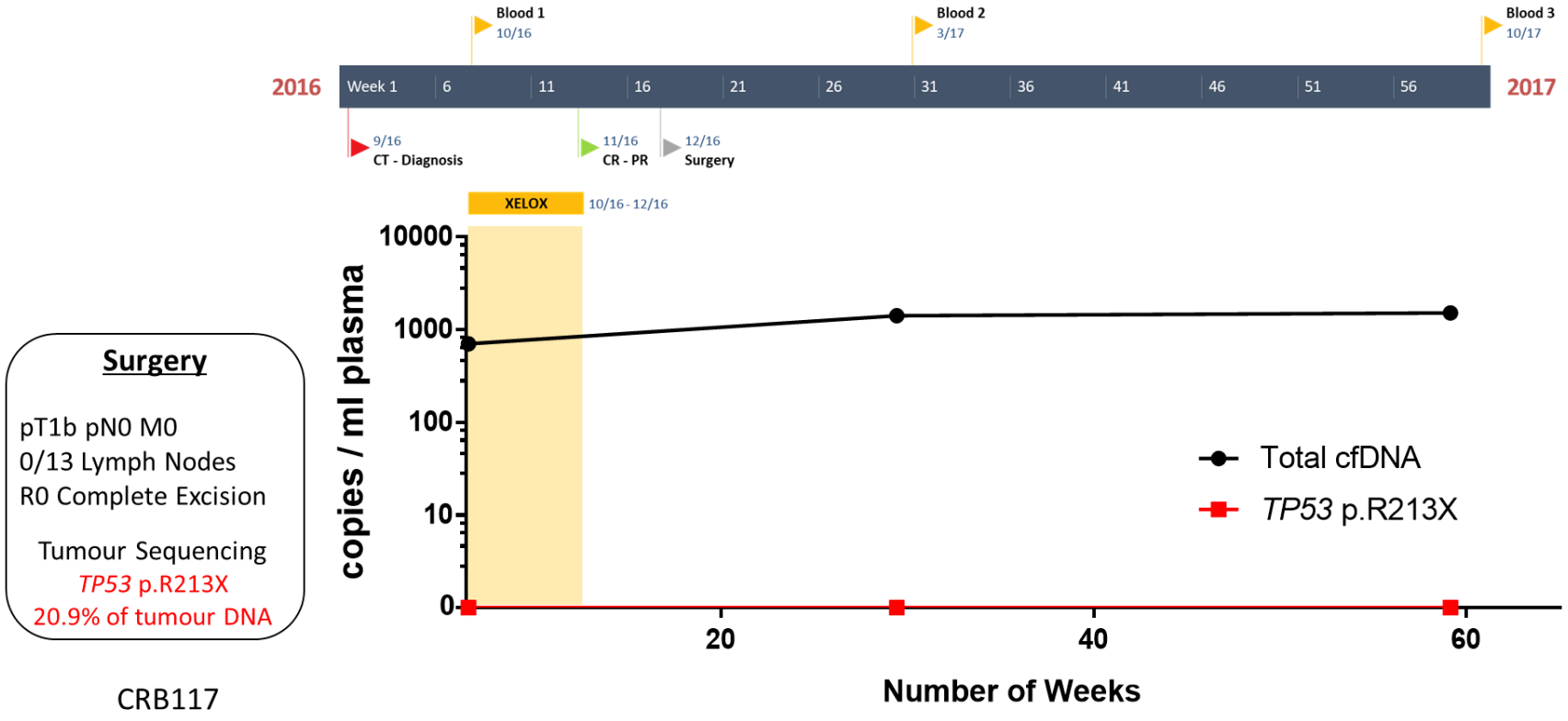
7.8.7 Patient CRB115



Blood number		1	2	3
cfDNA copies /ml		2930	3333	1717
TP53 p.H197R	AF (%)	0	0	0
	Mutant copies/ml	0	0	0

Figure 89 - **Timeline for patient CRB115.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.H179R mutation not detected in the total cfDNA by ddPCR (red). The RHOA mutation was not analysed by ddPCR as no assay was available.

7.8.8 Patient CRB117



Blood Number		1	3	4
cfDNA copies /ml		707	1414	1525
TP53 p.R213X	AF (%)	0	0	0
	Mutant copies/ml	0	0	0

Figure 90 - **Timeline for patient CRB117.** Showing clinical events in relation to total cfDNA copies/ml (black) and the *TP53* p.R213X mutation not detected in the total cfDNA by ddPCR (red). PR = Partial response.

7.8.9 ddPCR total cfDNA analysis for patients CRB144 and CRB145

Patient	Mutation	Tumour DNA mutation frequency	Total total cfDNA	ctDNA as represented by specific mutation	
			Copies/ml	VAF (%)	copies/ml
CRB144	*TP53* p.S215R	57.2%	2828	0	0
CRB145	*TP53* p.E286K	43.0%	2136	0	0

Table 50 – **Tumour and total cfDNA analysis for patients CRB144 and 145**.

7.9 Overview of tissue IHC and ddPCR results

Sample ID	HER2 status	RPPH1	CNTNAP1	GAPDH100
62 GT04.1	1+	3.76	1.72	2.98
66 GT07.1	1+	2.82	2.24	4.67
67 GT06.1	0	1.16	1.73	2.65
69 GT18.1	2+ FISH neg	1.97	1.64	2.73
75 GT17.1	0	1.49	1.7	1.57
83 GT25.1	2+ FISH neg	2	1.65	2.39
84 GT24.1	2+ FISH neg	2.63	1.95	2.25
89 GT32.1	3+	2.86	2.87	3.36
120 H520/17	3+	53.2	54.9	54.0
H521/17	3+	3.69	4.89	7.9
H523/17	2+ FISH pos	1.61	1.7	1.74
H525/17	3+	182	158	178
H526/17	3+	1.23	1.68	1.75
122 H527/17	3+	79	69	86
121 H528/17	2+ FISH pos	3.8	4.44	5.2
119 H529/17	2+ FISH pos	1.94	2.03	2.79
H577/17	2+ FISH pos	2.41	2.11	2.62
H578/17	3+	2.37	2.22	2.24
H579/17	?	144	118	178
H530/17	2+ FISH pos	1.92	2.17	4.59
118 H531/17	3+	30	19.7	41.8
SKBR3 +VE		17.3	33.4	17.9
gDNA -VE		2.1	1.97	2.25

Table 51 – **Overview of results of IHC and copy number as determined by ddPCR using 3 references.**

8 References

ABBOSH, C., BIRKBAK, N. J., WILSON, G. A., et al. 2017. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature,* 545**,** 446-451.

ABEL, H. J. & DUNCAVAGE, E. J. 2013. Detection of structural DNA variation from next generation sequencing data: a review of informatic approaches. *Cancer Genet,* 206**,** 432-40.

ABREU VELEZ, A. M. & HOWARD, M. S. 2010. Diagnosis and treatment of cutaneous paraneoplastic disorders. *Dermatol Ther,* 23**,** 662-75.

AGRAWAL, N., JIAO, Y., BETTEGOWDA, C., et al. 2012. Comparative genomic analysis of esophageal adenocarcinoma and squamous cell carcinoma. *Cancer Discov,* 2**,** 899-905.

AHLBORN, L. B., ROHRBERG, K. S., GABRIELAITE, M., et al. 2019. Application of cell-free DNA for genomic tumor profiling: a feasibility study. *Oncotarget,* 10**,** 1388-1398.

AKIBA, S., KORIYAMA, C., HERRERA-GOEPFERT, R., et al. 2008. Epstein-Barr virus associated gastric carcinoma: epidemiological and clinicopathological features. *Cancer Sci,* 99**,** 195-201.

AL-BATRAN, S.-E., HOMANN, N., SCHMALENBERG, H., et al. 2017a. Perioperative chemotherapy with docetaxel, oxaliplatin, and fluorouracil/leucovorin (FLOT) versus epirubicin, cisplatin, and fluorouracil or capecitabine (ECF/ECX) for resectable gastric or gastroesophageal junction (GEJ) adenocarcinoma (FLOT4-AIO): A multicenter, randomized phase 3 trial. *Journal of Clinical Oncology,* 35**,** 4004-4004.

AL-BATRAN, S. E., HOMANN, N., PAULIGK, C., et al. 2017b. Effect of Neoadjuvant Chemotherapy Followed by Surgical Resection on Survival in Patients With Limited Metastatic Gastric or Gastroesophageal Junction Cancer: The AIO-FLOT3 Trial. *JAMA Oncol,* 3**,** 1237-1244.

ALEXANDROV, L. B., NIK-ZAINAL, S., WEDGE, D. C., et al. 2013. Signatures of mutational processes in human cancer. *Nature,* 500**,** 415-21.

ALLOTT, E. H., GERADTS, J., SUN, X., et al. 2016. Intratumoral heterogeneity as a source of discordance in breast cancer biomarker classification. *Breast Cancer Res,* 18**,** 68-79.

ANKER, P., STROUN, M. & MAURICE, P. A. 1975. Spontaneous release of DNA by human blood lymphocytes as shown in an in vitro system. *Cancer Res,* 35**,** 2375-82.

ANSELL, S. M., LESOKHIN, A. M., BORRELLO, I., et al. 2015. PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N Engl J Med,* 372**,** 311-9.

APARICIO, S. & CALDAS, C. 2013. The implications of clonal genome evolution for cancer medicine. *N Engl J Med,* 368**,** 842-51.

ASSOU, S., AIT-AHMED, O., EL MESSAOUDI, S., et al. 2014. Non-invasive pre-implantation genetic diagnosis of X-linked disorders. *Med Hypotheses,* 83**,** 506-8.

BAHASSI EL, M. & STAMBROOK, P. J. 2014. Next-generation sequencing technologies: breaking the sound barrier of human genetics. *Mutagenesis,* 29**,** 303-10.

BANERJI, S., CIBULSKIS, K., RANGEL-ESCARENO, C., et al. 2012. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature,* 486**,** 405-9.

BANG, Y. J., VAN CUTSEM, E., FEYEREISLOVA, A., et al. 2010. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet,* 376**,** 687-97.

BEIJE, N., SIEUWERTS, A. M., KRAAN, J., et al. 2018. Estrogen receptor mutations and splice variants determined in liquid biopsies from metastatic breast cancer patients. *Mol Oncol,* 12**,** 48-57.

BEITER, T., FRAGASSO, A., HUDEMANN, J., et al. 2011. Short-term treadmill running as a model for studying cell-free DNA kinetics in vivo. *Clin Chem,* 57**,** 633-6.

BERNAL, M., RUIZ-CABELLO, F., CONCHA, A., et al. 2012. Implication of the beta2-microglobulin gene in the generation of tumor escape phenotypes. *Cancer Immunol Immunother,* 61**,** 1359-71.

BEROUKHIM, R., MERMEL, C. H., PORTER, D., et al. 2010. The landscape of somatic copy-number alteration across human cancers. *Nature,* 463**,** 899-905.

BETTEGOWDA, C., SAUSEN, M., LEARY, R. J., et al. 2014. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med,* 6**,** 224ra24.

BLAIR, V., MARTIN, I., SHAW, D., et al. 2006. Hereditary diffuse gastric cancer: diagnosis and management. *Clin Gastroenterol Hepatol,* 4**,** 262-75.

BOKU, N. 2014. HER2-positive gastric cancer. *Gastric Cancer,* 17**,** 1-12.

BOLGER, A. M., LOHSE, M. & USADEL, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics,* 30**,** 2114-20.

BOSETTI, C., BERTUCCIO, P., LEVI, F., et al. 2012. The decline in breast cancer mortality in Europe: an update (to 2009). *Breast,* 21**,** 77-82.

BRIERLEY, J. D., GOSPODAROWICZ, M. K. & WITTEKIND, C. 2016. *TNM classification of malignant tumors. 8.,* Oxford, Wiley-Blackwell.

BROAD, I. 2018. *Picard Tools* [Online]. Available: http://broadinstitute.github.io/picard/ [Accessed 01/05/2018 2018].

BUROTTO, M., CHIOU, V. L., LEE, J. M., et al. 2014. The MAPK pathway across different malignancies: a new perspective. *Cancer,* 120**,** 3446-56.

BUTLER, T. M., JOHNSON-CAMACHO, K., PETO, M., et al. 2015. Exome Sequencing of Cell-Free DNA from Metastatic Cancer Patients Identifies Clinically Actionable Mutations Distinct from Primary Disease. *PLoS One,* 10**,** e0136407.

CANCER GENOME ATLAS RESEARCH, N. 2014. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature,* 513**,** 202-9.

CANCER GENOME ATLAS RESEARCH, N., ANALYSIS WORKING GROUP: ASAN, U., AGENCY, B. C. C., et al. 2017. Integrated genomic characterization of oesophageal carcinoma. *Nature,* 541**,** 169-175.

CARDOSO, F., COSTA, A., NORTON, L., et al. 2014. ESO-ESMO 2nd international consensus guidelines for advanced breast cancer (ABC2)dagger. *Ann Oncol,* 25**,** 1871-88.

CARDOSO, F., SENKUS, E., COSTA, A., et al. 2018. 4th ESO-ESMO International Consensus Guidelines for Advanced Breast Cancer (ABC 4). *Ann Oncol***,** 1-24.

CARLSON, R. W. 2005. The history and mechanism of action of fulvestrant. *Clin Breast Cancer,* 6 Suppl 1**,** S5-8.

CASTRO, C., BOSETTI, C., MALVEZZI, M., et al. 2014. Patterns and trends in esophageal cancer mortality and incidence in Europe (1980-2011) and predictions to 2015. *Ann Oncol,* 25**,** 283-90.

CHAN, K. C., JIANG, P., ZHENG, Y. W., et al. 2013. Cancer genome scanning in plasma: detection of tumor-associated copy number aberrations, single-nucleotide variants, and tumoral heterogeneity by massively parallel sequencing. *Clin Chem,* 59**,** 211-24.

CHANDRASOMA, P. T., DER, R., DALTON, P., et al. 2001. Distribution and significance of epithelial types in columnar-lined esophagus. *Am J Surg Pathol,* 25**,** 1188-93.

CHEN, H., LIU, H. & QING, G. 2018. Targeting oncogenic Myc as a strategy for cancer treatment. *Signal Transduct Target Ther,* 3**,** 1-7.

CHEN, X., BONNEFOI, H., DIEBOLD-BERGER, S., et al. 1999. Detecting tumor-related alterations in plasma or serum DNA of patients diagnosed with breast cancer. *Clin Cancer Res,* 5**,** 2297-303.

CHEN, Y., SHI, L., ZHANG, L., et al. 2008. The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. *J Biol Chem,* 283**,** 17969-78.

CHO, E. Y., PARK, K., DO, I., et al. 2013. Heterogeneity of ERBB2 in gastric carcinomas: a study of tissue microarray and matched primary and metastatic carcinomas. *Mod Pathol,* 26**,** 677-84.

CHUNG, S. M., KAO, J., HYJEK, E., et al. 2007. p53 in esophageal adenocarcinoma: a critical reassessment of mutation frequency and identification of 72Arg as the dominant allele. *Int J Oncol,* 31**,** 1351-5.

COHEN, J. D., LI, L., WANG, Y., et al. 2018. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science,* 359**,** 926-930.

COLOMBAT, M., FORT, M. P., CHOLLET, C., et al. 2006. Molecular remission in chronic myeloid leukemia patients with sustained complete cytogenetic remission after imatinib mesylate treatment. *Haematologica,* 91**,** 162-8.

CRUK. 2018a. *A trial looking at whether ctDNA can help doctors know who is at risk of breast cancer coming back and if pembrolizumab reduces ctDNA (c-TRAK TN)* [Online]. Available: https://www.cancerresearchuk.org/about-cancer/find-a-clinical-trial/a-trial-looking-whether-ctdna-can-help-doctors-know-who-risk-breast-cancer-coming-back-pembroluzimab-reduces-ctdna-ctraktn#undefined [Accessed 01/11/2018].

CRUK. 2018b. *A trial using a blood test to find certain gene changes and decide treatment for advanced breast cancer (plasmaMATCH)* [Online]. https://www.cancerresearchuk.org/: CRUK. Available: https://www.cancerresearchuk.org/about-cancer/find-a-clinical-trial/a-trial-using-a-blood-test-to-find-certain-gene-changes-and-decide-treatment-for-advanced-breast#undefined [Accessed 11/11/2018 2018].

CUNNINGHAM, D., ALLUM, W. H., STENNING, S. P., et al. 2006. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N Engl J Med,* 355**,** 11-20.

CUNNINGHAM, D., STARLING, N., RAO, S., et al. 2008. Capecitabine and oxaliplatin for advanced esophagogastric cancer. *N Engl J Med,* 358**,** 36-46.

CURTIT, E., NERICH, V., MANSI, L., et al. 2013. Discordances in estrogen receptor status, progesterone receptor status, and HER2 status between primary breast cancer and metastasis. *Oncologist,* 18**,** 667-74.

DAWSON, S. J., TSUI, D. W., MURTAZA, M., et al. 2013. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N Engl J Med,* 368**,** 1199-209.

DE BRUIN, E. C., MCGRANAHAN, N., MITTER, R., et al. 2014. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science,* 346**,** 251-6.

DE MATTOS-ARRUDA, L., MAYOR, R., NG, C. K., et al. 2015a. Cerebrospinal fluid-derived circulating tumour DNA better represents the genomic alterations of brain tumours than plasma. *Nat Commun,* 6**,** 8839-45.

DE MATTOS-ARRUDA, L., MAYOR, R., NG, C. K., et al. 2015b. Cerebrospinal fluid-derived circulating tumour DNA better represents the genomic alterations of brain tumours than plasma. *Nat Commun,* 6**,** 8839.

DE MATTOS-ARRUDA, L., WEIGELT, B., CORTES, J., et al. 2014. Capturing intra-tumor genetic heterogeneity by de novo mutation profiling of circulating cell-free tumor DNA: a proof-of-principle. *Ann Oncol,* 25**,** 1729-35.

DESAI, T. K., KRISHNAN, K., SAMALA, N., et al. 2012. The incidence of oesophageal adenocarcinoma in non-dysplastic Barrett's oesophagus: a meta-analysis. *Gut,* 61**,** 970-6.

DEVESA, S. S. & FRAUMENI, J. F., JR. 1999. The rising incidence of gastric cardia cancer. *J Natl Cancer Inst,* 91**,** 747-9.

DIAZ, L. A., JR., WILLIAMS, R. T., WU, J., et al. 2012. The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature,* 486**,** 537-40.

DIEHL, F., SCHMIDT, K., CHOTI, M. A., et al. 2008. Circulating mutant DNA to assess tumor dynamics. *Nat Med,* 14**,** 985-90.

DO, H., WONG, S. Q., LI, J., et al. 2013. Reducing sequence artifacts in amplicon-based massively parallel sequencing of formalin-fixed paraffin-embedded DNA by enzymatic depletion of uracil-containing templates. *Clin Chem,* 59**,** 1376-83.

DOHM, J. C., LOTTAZ, C., BORODINA, T., et al. 2008. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Research,* 36**,** e105-e105.

DUFFY, M. J. 2015. Use of Biomarkers in Screening for Cancer. *Adv Exp Med Biol,* 867**,** 27-39.

DULAK, A. M., SCHUMACHER, S. E., VAN LIESHOUT, J., et al. 2012. Gastrointestinal Adenocarcinomas of the Esophagus, Stomach, and Colon Exhibit Distinct Patterns of Genome Instability and Oncogenesis. *Cancer Research,* 72**,** 4383-4393.

DULAK, A. M., STOJANOV, P., PENG, S., et al. 2013. Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity. *Nat Genet,* 45**,** 478-86.

ENGLAND, G. 2018. *The 100,000 Genomes Project* [Online]. Available: https://www.genomicsengland.co.uk/the-100000-genomes-project/ [Accessed 01/11/2018].

FAVERO, F., JOSHI, T., MARQUARD, A. M., et al. 2015. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann Oncol,* 26**,** 64-70.

FERLAY, J., SOERJOMATARAM, I., ERVIK, M., et al. 2013. *GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11.* [Online]. Lyon, France: International Agency for Research on Cancer. Available: http://globocan.iarc.fr [Accessed 11/01/2016 2016].

FINDLAY, J. M., CASTRO-GINER, F., MAKINO, S., et al. 2016. Differential clonal evolution in oesophageal cancers in response to neo-adjuvant chemotherapy. *Nat Commun,* 7**,** 11111-24.

FISHER, B., ANDERSON, S., BRYANT, J., et al. 2002. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. *N Engl J Med,* 347**,** 1233-41.

FORD, H. E. R., MARSHALL, A., BRIDGEWATER, J. A., et al. 2014. Docetaxel versus active symptom control for refractory oesophagogastric adenocarcinoma (COUGAR-02): an open-label, phase 3 randomised controlled trial. *The Lancet Oncology,* 15**,** 78-86.

FOSTER, J. M., OUMIE, A., TOGNERI, F. S., et al. 2015. Cross-laboratory validation of the OncoScan(R) FFPE Assay, a multiplex tool for whole genome tumour profiling. *BMC Med Genomics,* 8**,** 5-18.

FRENEL, J. S., CARREIRA, S., GOODALL, J., et al. 2015. Serial Next-Generation Sequencing of Circulating Cell-Free DNA Evaluating Tumor Clone Response To Molecularly Targeted Drug Administration. *Clin Cancer Res,* 21**,** 4586-96.

FRIBBENS, C., GARCIA MURILLAS, I., BEANEY, M., et al. 2018. Tracking evolution of aromatase inhibitor resistance with circulating tumour DNA analysis in metastatic breast cancer. *Ann Oncol,* 29**,** 145-153.

FRIBBENS, C., O'LEARY, B., KILBURN, L., et al. 2016. Plasma ESR1 Mutations and the Treatment of Estrogen Receptor-Positive Advanced Breast Cancer. *J Clin Oncol,* 34**,** 2961-8.

FURQAN, M., MUKHI, N., LEE, B., et al. 2013. Dysregulation of JAK-STAT pathway in hematological malignancies and JAK inhibitors for clinical application. *Biomark Res,* 1**,** 1-10.

GALE, D., LAWSON, A. R. J., HOWARTH, K., et al. 2018. Development of a highly sensitive liquid biopsy platform to detect clinically-relevant cancer mutations at low allele fractions in cell-free DNA. *PLoS One,* 13**,** e0194630.

GARCIA-MURILLAS, I., SCHIAVON, G., WEIGELT, B., et al. 2015. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci Transl Med,* 7**,** 302ra133.

GE, W., TENG, B. W., YU, D. C., et al. 2014. Dermatosis as the initial presentation of gastric cancer: two cases. *Chin J Cancer Res,* 26**,** 632-8.

GERLINGER, M., ROWAN, A. J., HORSWELL, S., et al. 2012. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med,* 366**,** 883-92.

GIROUX, V. & RUSTGI, A. K. 2017. Metaplasia: tissue injury adaptation and a precursor to the dysplasia-cancer sequence. *Nat Rev Cancer,* 17**,** 594-604.

GOH, J. Y., FENG, M., WANG, W., et al. 2017. Chromosome 1q21.3 amplification is a trackable biomarker and actionable target for breast cancer recurrence. *Nat Med,* 23**,** 1319-1330.

GRIFFITHS, A., GELBART, W. & MILLER, J. 1999. Chapter 7: Gene Mutations *Modern Genetic Analysis.* New York W. H. Freeman. 259-312.

GUTTERY, D. S., PAGE, K., HILLS, A., et al. 2015. Noninvasive detection of activating estrogen receptor 1 (ESR1) mutations in estrogen receptor-positive metastatic breast cancer. *Clin Chem,* 61**,** 974-82.

HAMAKAWA, T., KUKITA, Y., KUROKAWA, Y., et al. 2014. Monitoring gastric cancer progression with circulating tumour DNA. *British Journal of Cancer,* 112**,** 352-356.

HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *Cell,* 144**,** 646-74.

HARDENBOL, P., BANER, J., JAIN, M., et al. 2003. Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat Biotechnol,* 21**,** 673-8.

HARDIE, L. J., DARNTON, S. J., WALLIS, Y. L., et al. 2005. p16 expression in Barrett's esophagus and esophageal adenocarcinoma: association with genetic and epigenetic alterations. *Cancer Lett,* 217**,** 221-30.

HARLÉ, A., FILHINE-TRESARRIEU, P., HUSSON, M., et al. 2016. Rare RAS Mutations in Metastatic Colorectal Cancer Detected During Routine RAS Genotyping Using Next Generation Sequencing. *Targ Oncol,* 11**,** 363-370.

HARTGRINK, H. H., JANSEN, E. P., VAN GRIEKEN, N. C., et al. 2009. Gastric cancer. *Lancet,* 374**,** 477-90.

HAYAKAWA, Y., SETHI, N., SEPULVEDA, A. R., et al. 2016. Oesophageal adenocarcinoma and gastric cancer: should we mind the gap? *Nat Rev Cancer,* 16**,** 305-18.

HAYASHI, N. & YAMAUCHI, H. 2012. Role of circulating tumor cells and disseminated tumor cells in primary breast cancer. *Breast Cancer,* 19**,** 110-7.

HEHIR-KWA, J. Y., PFUNDT, R. & VELTMAN, J. A. 2015. Exome sequencing and whole genome sequencing for the detection of copy number variation. *Expert Rev Mol Diagn,* 15**,** 1023-32.

HERRERA, V. & PARSONNET, J. 2009. Helicobacter pylori and gastric adenocarcinoma. *Clin Microbiol Infect,* 15**,** 971-6.

HINDSON, B. J., NESS, K. D., MASQUELIER, D. A., et al. 2011. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem,* 83**,** 8604-10.

HODIS, E., WATSON, I. R., KRYUKOV, G. V., et al. 2012. A landscape of driver mutations in melanoma. *Cell,* 150**,** 251-63.

HOLLAND, P. M., ABRAMSON, R. D., WATSON, R., et al. 1991. Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proc Natl Acad Sci U S A,* 88**,** 7276-80.

HONG, L., HAN, Y., LIU, J., et al. 2013. Fibroblast growth factor receptor 2: a therapeutic target in gastric cancer. *Expert Rev Gastroenterol Hepatol,* 7**,** 759-65.

HSU, Y. H. & SHAW, C. K. 2000. The DCC protein expression in breast carcinoma. *Kaohsiung J Med Sci,* 16**,** 233-40.

HUDECOVA, I. 2015. Digital PCR analysis of circulating nucleic acids. *Clin Biochem,* 48**,** 948-56.

HUDIS, C. A. 2007. Trastuzumab--mechanism of action and use in clinical practice. *N Engl J Med,* 357**,** 39-51.

IAFRATE, A. J., FEUK, L., RIVERA, M. N., et al. 2004. Detection of large-scale variation in the human genome. *Nat Genet,* 36**,** 949-51.

ICR. 2018. *PlasmaMATCH* [Online]. https://www.icr.ac.uk/. Available: https://www.icr.ac.uk/our-research/centres-and-collaborations/centres-at-the-icr/clinical-trials-and-statistics-unit/clinical-trials/plasmamatch [Accessed 01/11/2018].

INC., G. 2018. *mskcc / vcf2maf* [Online]. Github Inc. . Available: https://github.com/mskcc/vcf2maf [Accessed 12/05/2018 2018].

JAHR, S., HENTZE, H., ENGLISCH, S., et al. 2001. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res,* 61**,** 1659-65.

JAMAL-HANJANI, M., WILSON, G. A., MCGRANAHAN, N., et al. 2017. Tracking the Evolution of Non-Small-Cell Lung Cancer. *N Engl J Med,* 376**,** 2109-2121.

JANKOWSKI, J. A., WRIGHT, N. A., MELTZER, S. J., et al. 1999. Molecular evolution of the metaplasia-dysplasia-adenocarcinoma sequence in the esophagus. *Am J Pathol,* 154**,** 965-73.

JATOI, I., CHEN, B. E., ANDERSON, W. F., et al. 2007. Breast cancer mortality trends in the United States according to estrogen receptor status and age at diagnosis. *J Clin Oncol,* 25**,** 1683-90.

JESELSOHN, R., YELENSKY, R., BUCHWALTER, G., et al. 2014. Emergence of constitutively active estrogen receptor-alpha mutations in pretreated advanced estrogen receptor-positive breast cancer. *Clin Cancer Res,* 20**,** 1757-67.

JOSHI, H. P., SUBRAMANIAN, I. V., SCHNETTLER, E. K., et al. 2014. Dynamin 2 along with microRNA-199a reciprocally regulate hypoxia-inducible factors and ovarian cancer metastasis. *Proc Natl Acad Sci U S A,* 111**,** 5331-6.

JUNG, K., FLEISCHHACKER, M. & RABIEN, A. 2010. Cell-free DNA in the blood as a solid tumor biomarker--a critical appraisal of the literature. *Clin Chim Acta,* 411**,** 1611-24.

KHIER, S. & LOHAN, L. 2018. Kinetics of circulating cell-free DNA for biomedical applications: critical appraisal of the literature. *Future Sci OA,* 4**,** Fso295.

KIM, K., SHIN, D. G., PARK, M. K., et al. 2014. Circulating cell-free DNA as a promising biomarker in patients with gastric cancer: diagnostic validity and significant reduction of cfDNA after surgical resection. *Ann Surg Treat Res,* 86**,** 136-42.

KIM, R., TAN, A., CHOI, M., et al. 2013. Geographic differences in approach to advanced gastric cancer: Is there a standard approach? *Crit Rev Oncol Hematol,* 88**,** 416-26.

KIM, S. T., LEE, J., HONG, M., et al. 2015. The NEXT-1 (Next generation pErsonalized tX with mulTi-omics and preclinical model) trial: prospective molecular screening trial of metastatic solid cancer patients, a feasibility analysis. *Oncotarget,* 6**,** 33358-68.

KINDE, I., WU, J., PAPADOPOULOS, N., et al. 2011. Detection and quantification of rare mutations with massively parallel sequencing. *Proc Natl Acad Sci U S A,* 108**,** 9530-5.

KIVIOJA, T., VAHARAUTIO, A., KARLSSON, K., et al. 2011. Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods,* 9**,** 72-4.

KOREN, R., DEKEL, Y., SHERMAN, E., et al. 2003. The expression of DCC protein in female breast cancer. *Breast Cancer Res Treat,* 80**,** 215-20.

KUTYAVIN, I. V., AFONINA, I. A., MILLS, A., et al. 2000. 3′-Minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. *Nucleic Acids Res,* 28**,** 655-61.

KWONG, Y. L., YEUNG, D. Y. & CHAN, J. C. 2009. Intrathecal chemotherapy for hematologic malignancies: drugs and toxicities. *Ann Hematol,* 88**,** 193-201.

LAGERGREN, J. & LAGERGREN, P. 2013. Recent developments in esophageal adenocarcinoma. *CA Cancer J Clin,* 63**,** 232-48.

LANMAN, R. B., MORTIMER, S. A., ZILL, O. A., et al. 2015. Analytical and Clinical Validation of a Digital Sequencing Panel for Quantitative, Highly Accurate Evaluation of Cell-Free Circulating Tumor DNA. *PLoS One,* 10**,** e0140712.

LEON, S. A., SHAPIRO, B., SKLAROFF, D. M., et al. 1977. Free DNA in the serum of cancer patients and the effect of therapy. *Cancer Res,* 37**,** 646-50.

LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics,* 25**,** 1754-60.

LI, J., DITTMAR, R. L., XIA, S., et al. 2017. Cell-free DNA copy number variations in plasma from colorectal cancer patients. *Mol Oncol,* 11**,** 1099-1111.

LI, J. Y., HO, J. C. & WONG, K. H. 2018a. T790M mutant copy number quantified via ddPCR predicts outcome after osimertinib treatment in lung cancer. *Oncotarget,* 9**,** 27929-27939.

LI, Z., LEVINE, K. M., BAHREINI, A., et al. 2018b. Upregulation of IRS1 Enhances IGF1 Response in Y537S and D538G ESR1 Mutant Breast Cancer Cells. *Endocrinology,* 159**,** 285-296.

LIANG, X., VACHER, S., BOULAI, A., et al. 2018. Targeted next-generation sequencing identifies clinically relevant somatic mutations in a large cohort of inflammatory breast cancer. *Breast Cancer Res,* 20**,** 88-100.

LIH, C.-J., HARRINGTON, R. D., SIMS, D. J., et al. 2017. Analytical Validation of the Next-Generation Sequencing Assay for a Nationwide Signal-Finding Clinical Trial. *The Journal of Molecular Diagnostics,* 19**,** 313-327.

LIPS, E. H., MICHAUT, M., HOOGSTRAAT, M., et al. 2015. Next generation sequencing of triple negative breast cancer to find predictors for chemotherapy response. *Breast Cancer Res,* 17**,** 134.

LIPSON, E. J., VELCULESCU, V. E., PRITCHARD, T. S., et al. 2014. Circulating tumor DNA analysis as a real-time method for monitoring tumor burden in melanoma patients undergoing treatment with immune checkpoint blockade. *J Immunother Cancer,* 2**,** 42.

LO, Y. M., ZHANG, J., LEUNG, T. N., et al. 1999. Rapid clearance of fetal DNA from maternal plasma. *Am J Hum Genet,* 64**,** 218-24.

MAHONEY, K. M., RENNERT, P. D. & FREEMAN, G. J. 2015. Combination cancer immunotherapy and new immunomodulatory targets. *Nat Rev Drug Discov,* 14**,** 561-84.

MANDEL, P. & METAIS, P. 1948. Les acides nucleiques du plasma sanguin chez l'homme. *C R Acad Sci,* 142**,** 241-3.

MANIER, S., PARK, J., CAPELLETTI, M., et al. 2018. Whole-exome sequencing of cell-free DNA and circulating tumor cells in multiple myeloma. *Nat Commun,* 9**,** 1691.

MAYAKONDA, A., LIN, D. C., ASSENOV, Y., et al. 2018. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res,* 28**,** 1747-1756.

MCKENNA, A., HANNA, M., BANKS, E., et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res,* 20**,** 1297-303.

MCLAREN, W., GIL, L., HUNT, S. E., et al. 2016. The Ensembl Variant Effect Predictor. *Genome Biol,* 17**,** 122.

MENG, S., TRIPATHY, D., SHETE, S., et al. 2004. HER-2 gene amplification can be acquired as breast cancer progresses. *Proc Natl Acad Sci U S A,* 101**,** 9393-8.

MILLER, W. R., BARTLETT, J., BRODIE, A. M., et al. 2008. Aromatase inhibitors: are there differences between steroidal and nonsteroidal aromatase inhibitors and do they matter? *Oncologist,* 13**,** 829-37.

MOLPARIA, B., NICHANI, E. & TORKAMANI, A. 2017. Assessment of circulating copy number variant detection for cancer screening. *PLoS One,* 12**,** e0180647.

MOMTAZ, P., PENTSOVA, E., ABDEL-WAHAB, O., et al. 2016. Quantification of tumor-derived cell free DNA(cfDNA) by digital PCR (DigPCR) in cerebrospinal fluid of patients with BRAFV600 mutated malignancies. *Oncotarget,* 7**,** 85430-85436.

MOUGALIAN, S. S., HERNANDEZ, M., LEI, X., et al. 2015. Ten-Year Outcomes of Patients With Breast Cancer With Cytologically Confirmed Axillary Lymph Node Metastases and Pathologic Complete Response After Primary Systemic Chemotherapy. *JAMA Oncol*, 1-9.

MU, Z., KLINOWSKA, T., DONG, X., et al. 2014. AZD8931, an equipotent, reversible inhibitor of signaling by epidermal growth factor receptor (EGFR), HER2, and HER3: preclinical activity in HER2 non-amplified inflammatory breast cancer models. *J Exp Clin Cancer Res,* 33**,** 47.

MURTAZA, M., DAWSON, S. J., POGREBNIAK, K., et al. 2015. Multifocal clonal evolution characterized using circulating tumour DNA in a case of metastatic breast cancer. *Nat Commun,* 6**,** 8760.

MURUGAESU, N., WILSON, G. A., BIRKBAK, N. J., et al. 2015. Tracking the genomic evolution of esophageal adenocarcinoma through neoadjuvant chemotherapy. *Cancer Discov,* 5**,** 821-31.

NACHMAN, J. B., SATHER, H. N., SENSEL, M. G., et al. 1998. Augmented post-induction therapy for children with high-risk acute lymphoblastic leukemia and a slow response to initial therapy. *N Engl J Med,* 338**,** 1663-71.

NAGATA, S. 2000. Apoptotic DNA fragmentation. *Exp Cell Res,* 256**,** 12-8.

NEWMAN, A. M., BRATMAN, S. V., TO, J., et al. 2014. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med,* 20**,** 548-54.

NGAN, H. Y. S., SECKL, M. J., BERKOWITZ, R. S., et al. 2018. Update on the diagnosis and management of gestational trophoblastic disease. *Int J Gynaecol Obstet,* 143 Suppl 2**,** 79-85.

NI, J., WENG, L., LIU, Y., et al. 2017. Dynamic monitoring of EGFR mutations in circulating cell-free DNA for EGFR-mutant metastatic patients with lung cancer: Early detection of drug resistance and prognostic significance. *Oncol Lett,* 13**,** 4549-4557.

NIK-ZAINAL, S., DAVIES, H., STAAF, J., et al. 2016. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature,* 534**,** 47-54.

NIK-ZAINAL, S., VAN LOO, P., WEDGE, D. C., et al. 2012. The life history of 21 breast cancers. *Cell,* 149**,** 994-1007.

NONES, K., WADDELL, N., WAYTE, N., et al. 2014. Genomic catastrophes frequently arise in esophageal adenocarcinoma and drive tumorigenesis. *Nat Commun,* 5**,** 5224-45.

O'LEARY, B., HREBIEN, S., MORDEN, J. P., et al. 2018. Early circulating tumor DNA dynamics and clonal selection with palbociclib and fulvestrant for breast cancer. *Nat Commun,* 9**,** 896-906.

OCHENDUSZKO, S., PUSKULLUOGLU, M., KONOPKA, K., et al. 2015. Comparison of efficacy and safety of first-line palliative chemotherapy with EOX and mDCF regimens in patients with locally advanced inoperable or metastatic HER2-negative gastric or gastroesophageal junction adenocarcinoma: a randomized phase 3 trial. *Med Oncol,* 32**,** 242.

OKAMOTO, I., SAKAI, K., MORITA, S., et al. 2014. Multiplex genomic profiling of non-small cell lung cancers from the LETS phase III trial of first-line S-1/carboplatin versus paclitaxel/carboplatin: results of a West Japan Oncology Group study. *Oncotarget,* 5**,** 2293-304.

OLSSON, E., WINTER, C., GEORGE, A., et al. 2015. Serial monitoring of circulating tumor DNA in patients with primary breast cancer for detection of occult metastatic disease. *EMBO Mol Med,* 7**,** 1034-47.

OPENSHAW, M. R., HARVEY, R. A., SEBIRE, N. J., et al. 2016. Circulating Cell Free DNA in the Diagnosis of Trophoblastic Tumors. *EBioMedicine,* 29**,** 146-52.

OPENSHAW, M. R., RICHARDS, C. J., GUTTERY, D. S., et al. 2017. The genetics of gastroesophageal adenocarcinoma and the use of circulating cell free DNA for disease detection and monitoring. *Expert Rev Mol Diagn,* 17**,** 459-470.

OSHIRO, C., KAGARA, N., NAOI, Y., et al. 2015. PIK3CA mutations in serum DNA are predictive of recurrence in primary breast cancer patients. *Breast Cancer Res Treat,* 150**,** 299-307.

PAGE, K., GUTTERY, D. S., FERNANDEZ-GARCIA, D., et al. 2017. Next Generation Sequencing of Circulating Cell-Free DNA for Evaluating Mutations and Gene Amplification in Metastatic Breast Cancer. *Clin Chem,* 63**,** 532-541.

PAGE, K., GUTTERY, D. S., ZAHRA, N., et al. 2013. Influence of plasma processing on recovery and analysis of circulating nucleic acids. *PLoS One,* 8**,** e77963.

PAGE, K., HAVA, N., WARD, B., et al. 2011. Detection of HER2 amplification in circulating free DNA in patients with breast cancer. *Br J Cancer,* 104**,** 1342-8.

PAGE, K., POWLES, T., SLADE, M. J., et al. 2006. The importance of careful blood processing in isolation of cell-free DNA. *Ann N Y Acad Sci,* 1075**,** 313-7.

PARLA, J. S., IOSSIFOV, I., GRABILL, I., et al. 2011. A comparative analysis of exome capture. *Genome Biol,* 12**,** R97.

PAULSON, T. G., XU, L., SANCHEZ, C., et al. 2006. Neosquamous epithelium does not typically arise from Barrett's epithelium. *Clin Cancer Res,* 12**,** 1701-6.

PENG, Q., VIJAYA SATYA, R., LEWIS, M., et al. 2015. Reducing amplification artifacts in high multiplex amplicon sequencing by using molecular barcodes. *BMC Genomics,* 16**,** 589-601.

PETLJAK, M. & ALEXANDROV, L. B. 2016. Understanding mutagenesis through delineation of mutational signatures in human cancer. *Carcinogenesis,* 37**,** 531-40.

POLK, D. B. & PEEK, R. M., JR. 2010. Helicobacter pylori: gastric cancer and beyond. *Nat Rev Cancer,* 10**,** 403-14.

PRAT, A. & PEROU, C. M. 2011. Deconstructing the molecular portraits of breast cancer. *Mol Oncol,* 5**,** 5-23.

QUINLAN, A. R. & HALL, I. M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics,* 26**,** 841-2.

RAKHA, E. A., PINDER, S. E., BARTLETT, J. M., et al. 2015. Updated UK Recommendations for HER2 assessment in breast cancer. *J Clin Pathol,* 68**,** 93-9.

ROSS, J. S. & CRONIN, M. 2011. Whole cancer genome sequencing by next-generation methods. *Am J Clin Pathol,* 136**,** 527-39.

ROTHE, F., LAES, J. F., LAMBRECHTS, D., et al. 2014. Plasma circulating tumor DNA as an alternative to metastatic biopsies for mutational analysis in breast cancer. *Ann Oncol,* 25**,** 1959-65.

RUSSNES, H. G., LINGJAERDE, O. C., BORRESEN-DALE, A. L., et al. 2017. Breast Cancer Molecular Stratification: From Intrinsic Subtypes to Integrative Clusters. *Am J Pathol,* 187**,** 2152-2162.

RUSTIN, G. J. & BURG, M. E. V. D. 2009. A randomized trial in ovarian cancer (OC) of early treatment of relapse based on CA125 level alone versus delayed treatment based on

conventional clinical indicators (MRC OV05/EORTC 55955 trials). *Journal of Clinical Oncology,* 27**,** 1-15.

SAIKI, A. Y., CAENEPEEL, S., COSGROVE, E., et al. 2015. Identifying the determinants of response to MDM2 inhibition. *Oncotarget,* 6**,** 7701-12.

SANGIOVANNI, J. P., CHEN, J., GUPTA, A. S., et al. 2015. Netrin-1 - DCC Signaling Systems and Age-Related Macular Degeneration. *PLoS One,* 10**,** e0125548.

SCHIAVON, G., HREBIEN, S., GARCIA-MURILLAS, I., et al. 2015. Analysis of ESR1 mutation in circulating tumor DNA demonstrates evolution during therapy for metastatic breast cancer. *Sci Transl Med,* 7**,** 313ra182.

SEO, A. N., KWAK, Y., KIM, D. W., et al. 2014. HER2 status in colorectal cancer: its clinical significance and the relationship between HER2 gene amplification and expression. *PLoS One,* 9**,** e98528.

SHAH, S. P., MORIN, R. D., KHATTRA, J., et al. 2009. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature,* 461**,** 809-13.

SHAH, S. P., ROTH, A., GOYA, R., et al. 2012. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature,* 486**,** 395-9.

SHARMA, P. & MONTGOMERY, E. 2013. Gastrointestinal dysplasia. *Pathology,* 45**,** 273-85.

SHAW, J. A., GUTTERY, D. S., HILLS, A., et al. 2017. Mutation analysis of cell-free DNA and single circulating tumor cells in metastatic breast cancer patients with high CTC counts. *Clin Cancer Res,* 23**,** 88-96.

SHAW, J. A., PAGE, K., BLIGHE, K., et al. 2012. Genomic analysis of circulating cell-free DNA infers breast cancer dormancy. *Genome Res,* 22**,** 220-31.

SHAW, J. A. & STEBBING, J. 2014. Circulating free DNA in the management of breast cancer. *Ann Transl Med,* 2**,** 3.

SHODA, K., ICHIKAWA, D., FUJITA, Y., et al. 2016. Monitoring the HER2 copy number status in circulating tumor DNA by droplet digital PCR in patients with gastric cancer. *Gastric Cancer*.

SHODA, K., ICHIKAWA, D., FUJITA, Y., et al. 2017. Monitoring the HER2 copy number status in circulating tumor DNA by droplet digital PCR in patients with gastric cancer. *Gastric Cancer,* 20**,** 126-135.

SHODA, K., MASUDA, K., ICHIKAWA, D., et al. 2014. HER2 amplification detected in the circulating DNA of patients with gastric cancer: a retrospective pilot study. *Gastric Cancer,* 18**,** 698-710.

SIEWERT, J. R. & STEIN, H. J. 1998. Classification of adenocarcinoma of the oesophagogastric junction. *Br J Surg,* 85**,** 1457-9.

SILVA, J. M., GARCIA, J. M., DOMINGUEZ, G., et al. 2002. Persistence of tumor DNA in plasma of breast cancer patients after mastectomy. *Ann Surg Oncol,* 9**,** 71-6.

SINGLA, H., LUDHIADCH, A., KAUR, R. P., et al. 2017. Recent advances in HER2 positive breast cancer epigenetics: Susceptibility and therapeutic strategies. *Eur J Med Chem,* 142**,** 316-327.

SORENSON, G. D., PRIBISH, D. M., VALONE, F. H., et al. 1994. Soluble normal and mutated DNA sequences from single-copy genes in human blood. *Cancer Epidemiol Biomarkers Prev,* 3**,** 67-71.

SPARANO, J. A., GRAY, R. J., MAKOWER, D. F., et al. 2018. Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer. *N Engl J Med,* 379**,** 111-121.

STASIK, S., SCHUSTER, C., ORTLEPP, C., et al. 2018. An optimized targeted Next-Generation Sequencing approach for sensitive detection of single nucleotide variants. *Biomolecular Detection and Quantification,* 15**,** 6-12.

STEPHENS, P. J., TARPEY, P. S., DAVIES, H., et al. 2012. The landscape of cancer genes and mutational processes in breast cancer. *Nature,* 486**,** 400-4.

STORCHOVA, Z. & KLOOSTERMAN, W. P. 2016. The genomic characteristics and cellular origin of chromothripsis. *Curr Opin Cell Biol,* 40**,** 106-113.

STROUN, M., LYAUTEY, J., LEDERREY, C., et al. 2001. About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. *Clin Chim Acta,* 313**,** 139-42.

SUGANO, K. 2015. Screening of gastric cancer in Asia. *Best Practice & Research Clinical Gastroenterology,* 29**,** 895-905.

SWANTON, C. 2012. Intratumor heterogeneity: evolution through space and time. *Cancer Res,* 72**,** 4875-82.

TALSMA, K., VAN HAGEN, P., GROTENHUIS, B. A., et al. 2012. Comparison of the 6th and 7th Editions of the UICC-AJCC TNM Classification for Esophageal Cancer. *Ann Surg Oncol,* 19**,** 2142-8.

TAMBORERO, D., RUBIO-PEREZ, C., DEU-PONS, J., et al. 2018. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med,* 10**,** 25-32.

TERWILLIGER, T. & ABDUL-HAY, M. 2017. Acute lymphoblastic leukemia: a comprehensive review and 2017 update. *Blood Cancer J,* 7**,** e577.

THIERRY, A. R., EL MESSAOUDI, S., GAHAN, P. B., et al. 2016. Origins, structures, and functions of circulating DNA in oncology. *Cancer Metastasis Rev,* 35**,** 347-76.

THIERRY, A. R., MOULIERE, F., EL MESSAOUDI, S., et al. 2014. Clinical validation of the detection of KRAS and BRAF mutations from circulating tumor DNA. *Nat Med,* 20**,** 430-5.

THOMA, C. 2018. Prostate cancer: Pitfalls of liquid biopsy tests. *Nat Rev Urol,* 15**,** 69.

THRIFT, A. P. 2016. The epidemic of oesophageal carcinoma: Where are we now? *Cancer Epidemiol,* 41**,** 88-95.

TOLEDO, R. A., GARRALDA, E., MITSI, M., et al. 2018. Exome Sequencing of Plasma DNA Portrays the Mutation Landscape of Colorectal Cancer and Discovers Mutated VEGFR2 Receptors as Modulators of Antiangiogenic Therapies. *Clin Cancer Res,* 24**,** 3550-3559.

TORGA, G. & PIENTA, K. J. 2018. Patient-Paired Sample Congruence Between 2 Commercial Liquid Biopsy Tests. *JAMA Oncol,* 4**,** 868-870.

VAN HAGEN, P., HULSHOF, M. C., VAN LANSCHOT, J. J., et al. 2012. Preoperative chemoradiotherapy for esophageal or junctional cancer. *N Engl J Med,* 366**,** 2074-84.

VARLEY, J. M., ATTWOOLL, C., WHITE, G., et al. 2001. Characterization of germline TP53 splicing mutations and their genetic and functional analysis. *Oncogene,* 20**,** 2647-54.

VELDMAN-JONES, M. H., BRANT, R., ROONEY, C., et al. 2015. Evaluating Robustness and Sensitivity of the NanoString Technologies nCounter Platform to Enable Multiplexed Gene Expression Analysis of Clinical Samples. *Cancer Res,* 75**,** 2587-93.

VERMA, S., MILES, D., GIANNI, L., et al. 2012. Trastuzumab emtansine for HER2-positive advanced breast cancer. *N Engl J Med,* 367**,** 1783-91.

VOGEL, C. L., COBLEIGH, M. A., TRIPATHY, D., et al. 2001. First-line Herceptin monotherapy in metastatic breast cancer. *Oncology,* 61 Suppl 2**,** 37-42.

VOLLBRECHT, C., LEHMANN, A., LENZE, D., et al. 2018. Validation and comparison of two NGS assays for the detection of EGFR T790M resistance mutation in liquid biopsies of NSCLC patients. *Oncotarget,* 9**,** 18529-18539.

WAGNER, A. D., UNVERZAGT, S., GROTHE, W., et al. 2010. Chemotherapy for advanced gastric cancer. *Cochrane Database Syst Rev***,** Cd004064.

WAN, L., PANTEL, K. & KANG, Y. 2013. Tumor metastasis: moving new biological insights into the clinic. *Nat Med,* 19**,** 1450-64.

WANG, X., LI, X., CHENG, Y., et al. 2015a. Copy number alterations detected by whole-exome and whole-genome sequencing of esophageal adenocarcinoma. *Hum Genomics,* 9**,** 22-37.

WANG, Y., SPRINGER, S., ZHANG, M., et al. 2015b. Detection of tumor-derived DNA in cerebrospinal fluid of patients with primary tumors of the brain and spinal cord. *Proceedings of the National Academy of Sciences of the United States of America,* 112**,** 9704-9709.

WANG, Z. C., LIN, M., WEI, L. J., et al. 2004. Loss of heterozygosity and its correlation with expression profiles in subclasses of invasive breast cancers. *Cancer Res,* 64**,** 64-71.

WEAVER, J. M., ROSS-INNES, C. S., SHANNON, N., et al. 2014. Ordering of mutations in preinvasive disease stages of esophageal carcinogenesis. *Nat Genet,* 46**,** 837-43.

WEBB, A., CUNNINGHAM, D., SCARFFE, J. H., et al. 1997. Randomized trial comparing epirubicin, cisplatin, and fluorouracil versus fluorouracil, doxorubicin, and methotrexate in advanced esophagogastric cancer. *J Clin Oncol,* 15**,** 261-7.

WEIGELT, B., PETERSE, J. L. & VAN 'T VEER, L. J. 2005. Breast cancer metastasis: markers and models. *Nat Rev Cancer,* 5**,** 591-602.

WILLIAMS, C., PONTEN, F., MOBERG, C., et al. 1999. A high frequency of sequence alterations is due to formalin fixation of archival specimens. *Am J Pathol,* 155**,** 1467-71.

WUEBBEN, E. L. & RIZZINO, A. 2017. The dark side of SOX2: cancer - a comprehensive overview. *Oncotarget,* 8**,** 44917-44943.

XU, B., TENG, L. H., SILVA, S. D., et al. 2014. The significance of dynamin 2 expression for prostate cancer progression, prognostication, and therapeutic targeting. *Cancer Med,* 3**,** 14-24.

YAMAMOTO, K. N., HIROTA, K., TAKEDA, S., et al. 2014. Evolution of pre-existing versus acquired resistance to platinum drugs and PARP inhibitors in BRCA-associated cancers. *PLoS One,* 9**,** e105724.

ZHANG, D., ZHANG, P., YANG, P., et al. 2017. Downregulation of ATP1A1 promotes cancer development in renal cell carcinoma. *Clin Proteomics,* 14**,** 15-27.