Detecting and tracking circulating tumour DNA copy number profiles during first line chemotherapy in oesophagogastric adenocarcinoma

Michael Davidson[1]*, Louise J. Barber[2]*, Andrew Woolston[2], Catherine Cafferkey[1], Sonia Mansukhani[2], Beatrice Griffiths[2], Sing-Yu Moorcraft[1], Isma Rana[1], Ruwaida Begum[1], Ioannis Assiotis[2], Nik Matthews[2], Sheela Rao[1], David Watkins[1], Ian Chau[1], David Cunningham[1], Naureen Starling[1] and Marco Gerlinger[1,2]

Affiliations:

[1]Gastrointestinal and Lymphoma Unit, Royal Marsden NHS Foundation Trust, London and Sutton, United Kingdom.

[2]Centre for Evolution and Cancer, Institute of Cancer Research, London, United Kingdom.

*Contributed equally to this work

Corresponding Author:

Dr Marco Gerlinger, MD, FRCP

Translational Oncogenomics Laboratory

Centre for Evolution and Cancer

The Institute of Cancer Research

237 Fulham Road

London SW3 6JB, UK

email: marco.gerlinger@icr.ac.uk

Word count: 4272

Acknowledgements:

We would like to thank all patients participating in the FOrMAT clinical trial and the clinical trials team at the Royal Marsden Hospital. This project has received funding from a Wellcome Trust Strategic Grant (105104/Z/14/Z) to the ICR Centre for Evolution and Cancer, the National Institute for Health Research Biomedical Research Centre for Cancer at the ICR/RMH, a Clinician Scientist Fellowship from Cancer Research UK and from the

European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 820137).

Abstract

DNA somatic copy number aberrations (SCNAs) are key drivers in oesophago-gastric adenocarcinoma (OGA). Whether minimally invasive SCNA analysis of circulating tumour (ct)DNA can predict treatment outcomes and reveal how SCNAs evolve during chemotherapy is unknown. We investigated this by low-coverage whole genome sequencing (lcWGS) of ctDNA from 30 patients with advanced OGA prior to first-line chemotherapy and on progression. SCNA profiles were detectable pre-treatment in 23/30 (76.7%) patients. The presence of liver metastases, primary tumour in situ or of oesophageal or junctional tumour location predicted for a high ctDNA fraction. A low ctDNA concentration associated with significantly longer overall survival. Neither chromosomal instability metrics nor ploidy correlated with chemotherapy outcome. Chromosome 2q and 8p gains before treatment were associated with chemotherapy responses. lcWGS identified all amplifications found by prior targeted tumour tissue sequencing in cases with detectable ctDNA, as well as finding additional changes. SCNA profiles changed during chemotherapy, indicating that cancer cell populations evolved during treatment, however no recurrent SCNA changes were acquired at progression. Tracking the evolution of OGA cancer cell populations in ctDNA is feasible during chemotherapy. The observation of genetic evolution warrants investigation in larger series and with higher resolution techniques to reveal potential genetic predictors of response and drivers of chemotherapy resistance. The presence of liver metastasis is a potential biomarker for the selection of patients with high ctDNA content for such studies.

Introduction

Gastric and oesophageal cancers are a challenging health issue, representing the third and sixth leading causes of global cancer mortality respectively [1]. Advances have been made in the genetic characterisation and development of novel targeted agents for the adenocarcinoma histological subtype, however the outlook for advanced disease remains poor with median overall survival not extending beyond 12 months in the majority of trials [2]. Recent large-scale sequencing projects have improved insights into the genomic landscape of the disease. The 2014 Cancer Genome Atlas (TCGA) analysis described four different subtypes of gastric cancer, with the most common CIN subtype being characterised by chromosomal instability, aneuploidy and, in many cases, focal amplification of receptor tyrosine kinases. The genomes of these cancers harbour multiple DNA somatic copy number alterations (SCNAs), defined as deviations in the number of whole chromosomes, chromosome arms or fragments from the normal number of two copies per cell. With the exception of *p53* mutations, which occur in 70-80% of oesophagogastric adenocarcinomas (OGA) of the CIN subtype, mutations in cancer driver genes are relatively rare in these cancers and SCNAs are considered the predominant type of genetic driver alterations [3] [4]. Common SCNAs identified in CIN tumours in these landmark sequencing studies include amplifications of chromosomal regions harbouring genes encoding for receptor tyrosine kinases or their ligands such as *ERBB2*, *EGFR* and *VEGFA*; as well as those involved pathways regulating proliferation (*MYC*), and cell cycle (*CCNE1*, *CCND1* and *CDK6*). These SCNAs have been implicated as key and, in the case of *ERBB2*/HER2, clinically actionable drivers in OGA [5] [6].

The CIN subtype is common among gastric cancers arising proximally from the oesophagogastric junction or cardia [3] and in oesophageal adenocarcinomas [4]. The 'genomically stable' subtype is characterised by few SCNAs and associated with the diffuse histological subtype of gastric cancer that commonly arises more distally from the stomach body [3]. The incidence of non-cardia gastric adenocarcinomas is declining in Western populations, whilst that of junctional and oesophageal adenocarcinomas is increasing [7]. These tumours are predominantly of the CIN subtype, and thus detection of SCNAs, in particular the clinically and biologically relevant driver events within these complex profiles, are important for the ongoing development of new biomarkers and therapies.

SCNAs have traditionally been analysed through microarray-based techniques, although more recently improved sensitivity for SCNA detection has been achieved through exome or whole genome sequencing (WGS). However due to cost, long turnaround times and intensive bioinformatics analysis requirements, such large scale genomics analyses are

often not feasible. Low coverage WGS (lcWGS), using a coverage of only 0.1-0.5x (i.e. where only 10-50% of the genome is sequenced), has been shown to be sufficient for reliable detection of SCNAs, with recent data showing superior SCNA calling compared to older array hybridisation-based standards [8]. Crucially, lcWGS can also be applied to analyse tumour derived circulating free (cf)DNA extracted from the plasma of cancer patients [9]. Such liquid biopsies offer clear practical advantages over conventional biopsies, including the minimally invasive nature of sample acquisition, relative ease of standardisation of sampling protocols, and the ability to obtain repeated samples over time. The latter is of particular interest as changes in SCNA profiles over the course of treatment may shed light on response and resistance mechanisms to existing chemotherapy agents as well as to novel targeted agents and immunotherapies.

Intratumour heterogeneity is recognised as a major challenge in the delivery of effective molecular targeted treatment in OGA [10] [11]. Copy number variation of molecular targets, as assessed in both tumour and cfDNA, has been shown to impact on therapeutic targeting of *ERBB2*, *FGFR* and *EGFR,* with high level amplifications being associated with more favourable responses [12] [13] [14]. Application of targeted genomic sequencing to cfDNA analysis has been shown to allow the detection of mutations which are heterogeneous within OGA [15] [16]. Such liquid biopsy techniques may also facilitate tracking of genetic profile changes over time, but this has not been applied to OGAs undergoing systemic therapy.

We applied lcWGS to cfDNA from 30 patients with advanced OGA to investigate whether SCNA analysis can predict responses to first-line chemotherapy, and how these profiles may evolve during chemotherapy treatment.

Results

The clinical and pathological characteristics of the 30 included patients are summarized in Table 1. Extracted cfDNA concentrations from plasma samples taken at pre-treatment baseline ranged from 1.37 to 74.04 ng/mL with a median of 8.88 ng/mL. With a minimum input quantity of 5 ng for lcWGS, sufficient cfDNA was available from all 30 patients. Univariate analysis showed that the presence of the primary tumour in situ was associated with a significantly increased cfDNA concentration (Table 2, 9.66 vs 4.81 ng/mL; p=0.0027, Mann Whitney test). The cfDNA concentration was numerically higher in patients with liver metastases vs. those without liver metastases (10.09 vs 6.80 ng/mL; p=0.1306,

Mann Whitney test), but this was not significant. No other clinical or pathological parameters were associated with pre-treatment cfDNA concentration.

Table 1: Clinical characteristics of included patients.

Histopathological variable		
Number of Cases:		30
Anatomic site of primary:	Gastric OGJ/ oesophageal	6 (20%) 24 (80%)
Histological subtype:	Intestinal Diffuse	28 (93%) 2 (7%)
Clinical stage at presentation:	Locally advanced Metastatic	3 (10%) 27 (90%)
HER2 status*:	Positive Negative	6 (20%) 24 (80%)
First line chemotherapy:	Platinum/fluoropyrimidine doublet Doublet+ anthracycline Doublet+ trastuzumab	9 (30%) 15 (50%) 6 (20%)
Metastatic sites: Liver	Yes No	16 (53%) 14 (47%)
Peritoneal	Yes No	6 (20%) 24 (80%)
Lung	Yes No	8 (27%) 22 (73%)
Number of metastatic organ sites:	0 - 1 ≥ 2	22 (73%) 8 (27%)
Primary tumour in situ:	Yes No	23 (77%) 7 (23%)
CA19-9 secretor:	Yes No	15 (50%) 15 (50%)

*defined as HER2 IHC +++ on baseline diagnostic specimen from patient clinical records; OGJ- Oesophagogastric junction

Sequencing was performed with 100 bp single-end reads and a target of 12 million reads per sample. The ichorCNA bioinformatics package [17] was used to reconstruct copy number profiles from sequencing data and to estimate the fraction of cfDNA that was derived from tumour cells (henceforth denoted as circulating *tumour* (ct)DNA content). Based on ichorCNA analysis, 7/30 cases (23.3%) had ctDNA content of zero, leaving 23 cases (76.7%) in which SCNA analysis could be performed. The seven cases with zero tumour content included all three tumours that were only locally advanced rather than metastatic in this cohort (Cases 2, 152, 195). The other four (57.1%) cases with zero tumour content had metastatic disease involving only a single organ site (Cases 52, 66, 119, 144). The ctDNA content showed a poor correlation with the total cfDNA concentration in the plasma (Figure 1A, Pearson correlation r^2=0.2312), suggesting that the release of ctDNA from tumour cells

and the total amount of cfDNA, which is a mix of DNA from malignant and non-malignant cells, are largely independent from each other. The presence of the primary tumour in situ (9.1% vs 0% median ctDNA content, p=0.0046, Mann Whitney test) and the presence of liver metastases (18.0% vs 7.2% median ctDNA content, p=0.0043, Mann Whitney test) significantly correlated with higher ctDNA content (Table 2 and Figure 1B). A greater ctDNA content was also observed in oesophageal and junctional tumours compared to gastric tumours (9.3% vs 3.3% median ctDNA content, p=0.0103, Mann Whitney test).

Figure 1 (A) No correlation between cfDNA concentration and the tumour-derived cfDNA fraction in 30 plasma samples from patients with treatment naïve metastatic gastro-oesophageal cancers. (**B)** Correlation between selected clinical features and ctDNA fraction (line denotes median; p-value Mann Whitney test). (**C**) Kaplan Meier survival analyses of pre-treatment samples grouping by high/intermediate/low cfDNA yield ng/mL plasma, (**D**) ichorCNA ctDNA fraction, and (**E**) ctDNA concentration ng/mL plasma (p-values Log-rank (Mantel-Cox) test).

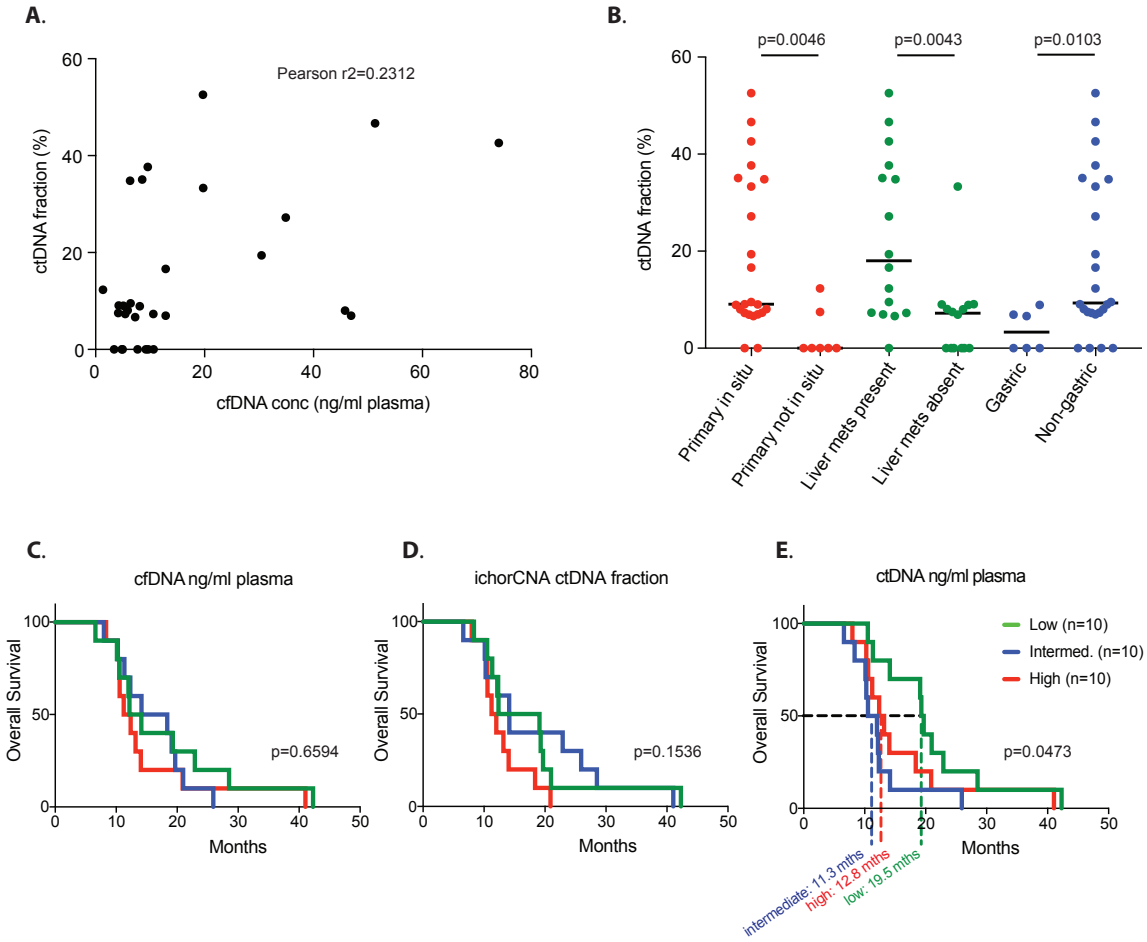


Table 2. Correlation of cfDNA concentration, median ichorCNA ctDNA fraction and ctDNA concentration with clinical and laboratory variables (p-values Mann-Whitney test).

Histopathological variable		N	Median cfDNA concentration (ng/mL plasma)	p-value	Median ctDNA fraction (%)	p-value	Median ctDNA concentration (ng/mL plasma)	p-value
Primary tumour in situ	Yes	23	9.66	0.0027	9.10	0.0046	2.14	<0.0001
	No	7	4.81		0.00		0.00	
Liver metastases present	Yes	16	10.09	0.1306	18.01	0.0043	2.18	0.0099
	No	14	6.80		7.23		0.35	
Primary tumour anatomic site	Gastric	6	8.65	0.8996	3.33	0.0103	0.24	0.1401
	Non-gastric	24	9.05		9.31		0.84	
No. of metastatic organ sites	0-1	22	8.31	0.5042	7.77	0.1528	0.47	0.9814
	≥2	8	1.22		14.47		0.58	
HER2 status	Positive	6	11.22	0.3739	8.81	0.4595	2.25	0.1713
	Negative	24	8.32		8.22		0.47	
CA19-9 secretion	Yes	15	9.21	0.9999	8.10	0.5640	0.61	0.7733
	No	15	8.54		9.02		0.78	

Taken together, copy number profiles could be analysed from cfDNA in 76.7% of cases and three distinct characteristics (primary tumour in situ, presence of liver metastases and oesophageal/junctional primary tumour location) associated with high ctDNA content, with liver metastases showing the highest tumour fraction of 18% (median).

We next investigated whether any pre-treatment cfDNA metrics correlate with overall survival (OS). Neither the total cfDNA concentration extracted from plasma (Figure 1C), nor the ctDNA content estimated by ichorCNA (Figure 1D) correlated with overall survival. However, the absolute ctDNA concentration in the plasma revealed a significant overall survival (OS) difference (Figure 1E). The third of patients with the lowest absolute ctDNA concentration (mean 0.09 ng/mL) had a median OS of 19.5 months whereas those with intermediate (mean 0.92 ng/mL) and high (mean 10.12 ng/mL) absolute ctDNA concentration had a median OS of 11.3 and 12.8 months, respectively.

We next investigated whether any specific copy number aberrations or chromosomal instability metrics correlated with subsequent responses to chemotherapy (Figure 2A and B). The frequency of copy number gains or losses in 13 responders (based on best radiological response assessment with serial CT scans during treatment) (Figure 2C) was compared to those in 10 non-responders who had stable or progressive disease as best response (Figure 2D). Frequency plots showed an overall similar appearance in both groups. However, several chromosomes showed alterations that were unique to the responders (Figure 2E) and not present in the non-responder group (Figure 2F). Gains of chromosomes 2q and 8p were the most frequent (>1/3 of cases) unique aberrations observed only among responders (Figure 2E). A minimal consistent region of 28 Mb on Chr2q encompassing 182 genes was observed in five of 13 cases (34, 63, 68, 134 and 207). These 2q gains were in four cases a single copy number gain relative to ploidy. A 7.5 Mb minimal consistent region on Chr8p encompassing 17 genes (Supplemental Table) was detected in six cases (34, 45, 68, 99, 143, 183), four of which as multiple copies above ploidy. Of the uniquely gained genes, *MCPH1* (microcephalin) is notable as a key regulator of DNA damage response and a repressor of human telomerase reverse transcriptase function [18] and gains of *MCPH1* have been implicated in increased platinum sensitivity in non-small cell lung cancer [19] (Figure 2G). Chr8p also harbours *GATA4* which is frequently gained or amplified in OGA [4] [20], but this was located outside the unique region as gains of *GATA4* were observed in both responders and non-responders (Figure 2G). Other uniquely altered regions were less frequent and hence difficult to assess (Figure 2E). In contrast, only a single loss of a 12 Mb minimal consistent region encompassing 117 genes on Chr1p in four cases (123, 126, 90 and 158) was unique to the non-responder group (Figure 2F).

Figure 2 (A) Integer copy number profiles (500 kb bins) for pre-treatment samples, grouped by subsequent response or (**B**) non-response to treatment. Red= gain, Blue= loss, Black= ploidy. **(C)** Frequency plots showing the number of cases that show segment gains (red) or losses (blue) in the responder and (**D**) non-responder groups. **(E)** Frequency plots showing segment gains and losses that are unique to the responder group or (**F**) non-responder group. (**G**) Frequency of gain (red) and loss (blue) segments of chromosome 8p in the responder group (top) and non-responder group (bottom). The most frequent region of unique 8p gain is indicated, bounded by dotted lines. The locations of *MCPH1* and *GATA4* are delineated with a blue dashed line. Two additional non-responder cases showed focal amplifications (orange) of *GATA4,* which were identified with the 50 kb bin method but not the 500 kb ichorCNA analysis.

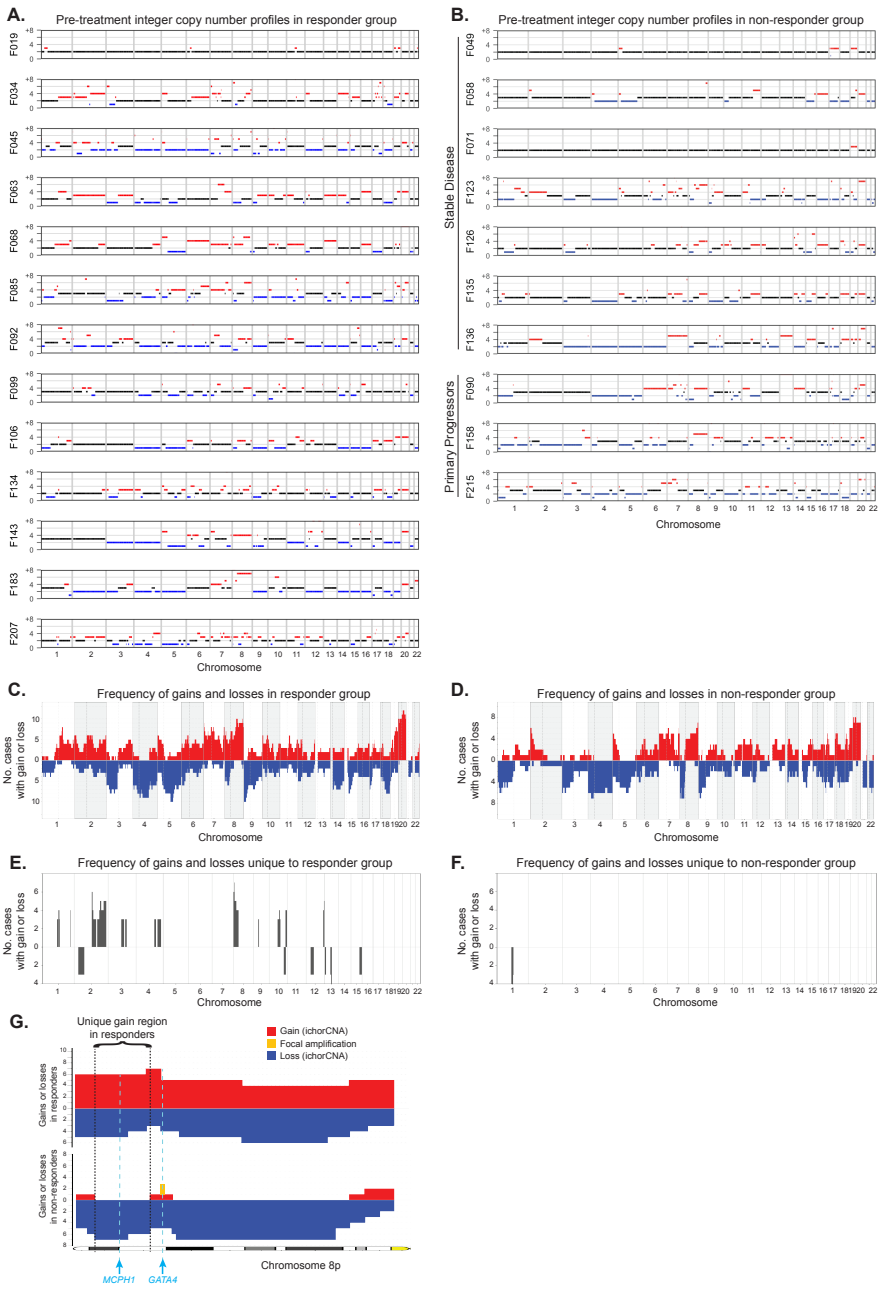
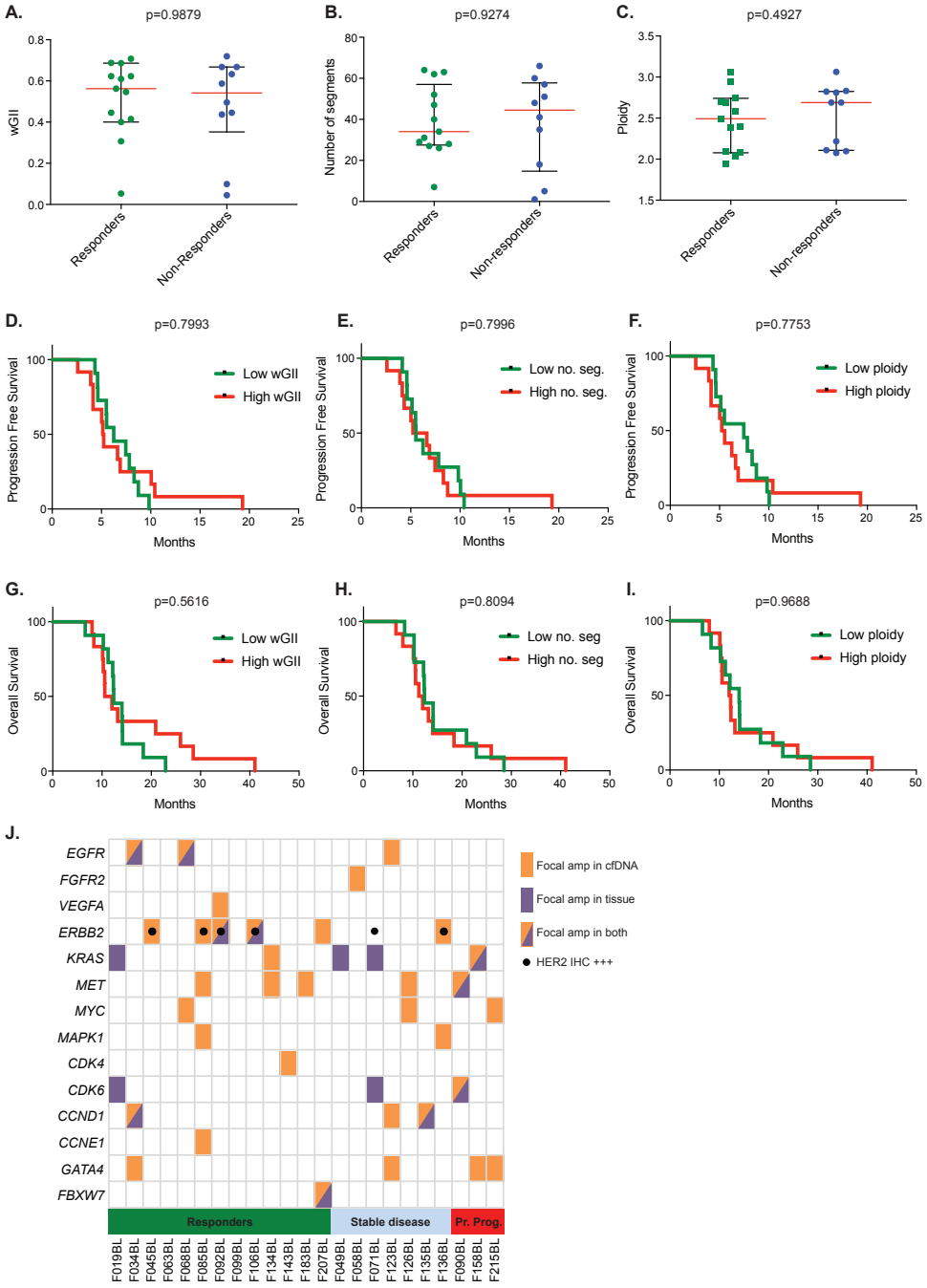


Figure 3 (**A**) Association of pre-treatment CIN metrics with subsequent treatment response by comparing analysis of genomic change relative to ploidy using weighted Genomic Instability index (wGII), (**B**) non-ploidy segment number, and (**C**) ploidy between responder

and non-responder groups (line denotes median and interquartile range; p-value Mann Whitney test). (**D**) Kaplan Meier progression free survival analyses grouping by high/low wGII, **(E)** non-ploidy segment number and (**F**) ploidy. **(G)** Kaplan Meier overall survival analyses grouping by high/low wGII, (**H**) non-ploidy segment number and (**I**) ploidy. (**J**) Heatmap showing focal gene amplifications (50 kb bins) detected by cfDNA lcWGS at pre-treatment (orange) or by archival target sequencing (purple) in each case. Black dots indicate cases classed as HER2+ by immunohistochemistry. Green= responder group, blue= stable group, red= primary progressor group.



Chromosomal instability (CIN) has been associated with poor outcomes and treatment responses in several cancer types [21] [22]. We hence assessed whether CIN-metrics including the weighted genomic instability index (wGII) [23] [24] (Figure 3A), the number of gained or lost chromosomal segments (Figure 3B) or ploidy (Figure 3C) associated with responses or could predict survival in our cohort. None of these metrics showed a significant difference in responders vs. non-responders or an association with progression-free (Figure 3D-F) or overall survival (Figure 3G-I). Taken together, the presence of Chr2q and 8p gains in pre-treatment ctDNA showed an association with chemotherapy responses. In contrast, we could not identify a role of CIN metrics to predict patient outcomes in OGA.

The ichorCNA analysis divides chromosomes into 500kb large bins to robustly assess the copy number state of these segments. Focal genomic amplifications are often narrow [4] (down to a few dozen kbps) and may have been overlooked as a consequence. Therefore, to further interrogate whether focal amplifications could be detected in the lcWGS data, we applied a 50kbp bin approach [25]. This revealed narrow high-level amplifications of several OGA driver genes [3] [4] (Figure 3J). Any of the high level amplifications (*EGFR, ERBB2, KRAS, MET, MYC, MAPK1/ERK2, CCND1* and *GATA4*) that were observed in two or more cases were detected in both responders and in non-responders. Several others were only observed once and were hence too rare to draw any conclusions. Thus, high-level amplifications detected pre-treatment did not associate with chemotherapy responses.

As part of the FOrMAT clinical trial, archival formalin-fixed paraffin-embedded diagnostic or resection samples were sequenced with a custom panel targeting 46 genes that had prognostic or predictive significance, or were potential targets in existing or upcoming clinical trials [26]. Amplifications of *EGFR*, *CCND1*, *CDK6*, *MET*, *ERBB2*, *KRAS*, and *FBXW7* had been identified in tissue samples from 11 cases (19, 34, 49, 68, 71, 90, 92, 106, 135, 158, 207). No amplifications were observed in nine cases and archival target sequencing failed in three cases (45, 58, 123). cfDNA lcWGS of pre-treatment plasma re-identified all of the gene amplifications found by archival tumour sequencing in eight cases (Figure 3J). Compared to tissue sequencing, ctDNA analysis could not detect *CDK6* and/or *KRAS* amplifications in three cases that had low ctDNA content (Case 19: 9.1%; Case 49: 7.3%; Case 71: 8.1%). Importantly, in seven cases, cfDNA lcWGS identified additional amplifications of genes that were included in the FOrMAT sequencing panel but for which no amplification was detected in the archival tissue analysis: Case 85 (*MET* and *ERBB2* amplification in plasma), Case 126 (*MET*), Case 134 (*MET, KRAS*), Case 136 (*ERBB2*), Case 143 (*CDK4*), Case 183 (*MET*) and Case 207 (*ERBB2*). In addition, cfDNA sequencing

identified 11 amplifications (in nine cases) of genes that were not covered by the FOrMAT panel, including *GATA4*, *VEGFA* and *MYC*.

Of six cases (45, 71, 85, 92, 106, 136) that had been classified as HER2 positive based on standard immunohistochemistry testing of archival tissue, cfDNA sequencing detected *ERBB2* amplifications in five cases. Archival tissue sequencing had identified *ERBB2* amplifications in only two of five successfully sequenced cases (Figure 3J). In one case (71) immunohistochemical (IHC) analysis of archival tissue had identified HER2 positivity but no amplification was detected by either archival tumour sequencing or cfDNA lcWGS. Three of the *ERBB2* amplified cases (85, 92, 136) had concurrent amplifications in *MAPK1*, *MET*, or *VEGFA* in the cfDNA (Figure 3J).

Table 3. Comparison of ichorCNA estimated ctDNA fraction at pre-treatment and progression of first line chemotherapy (p-values Mann-Whitney test).

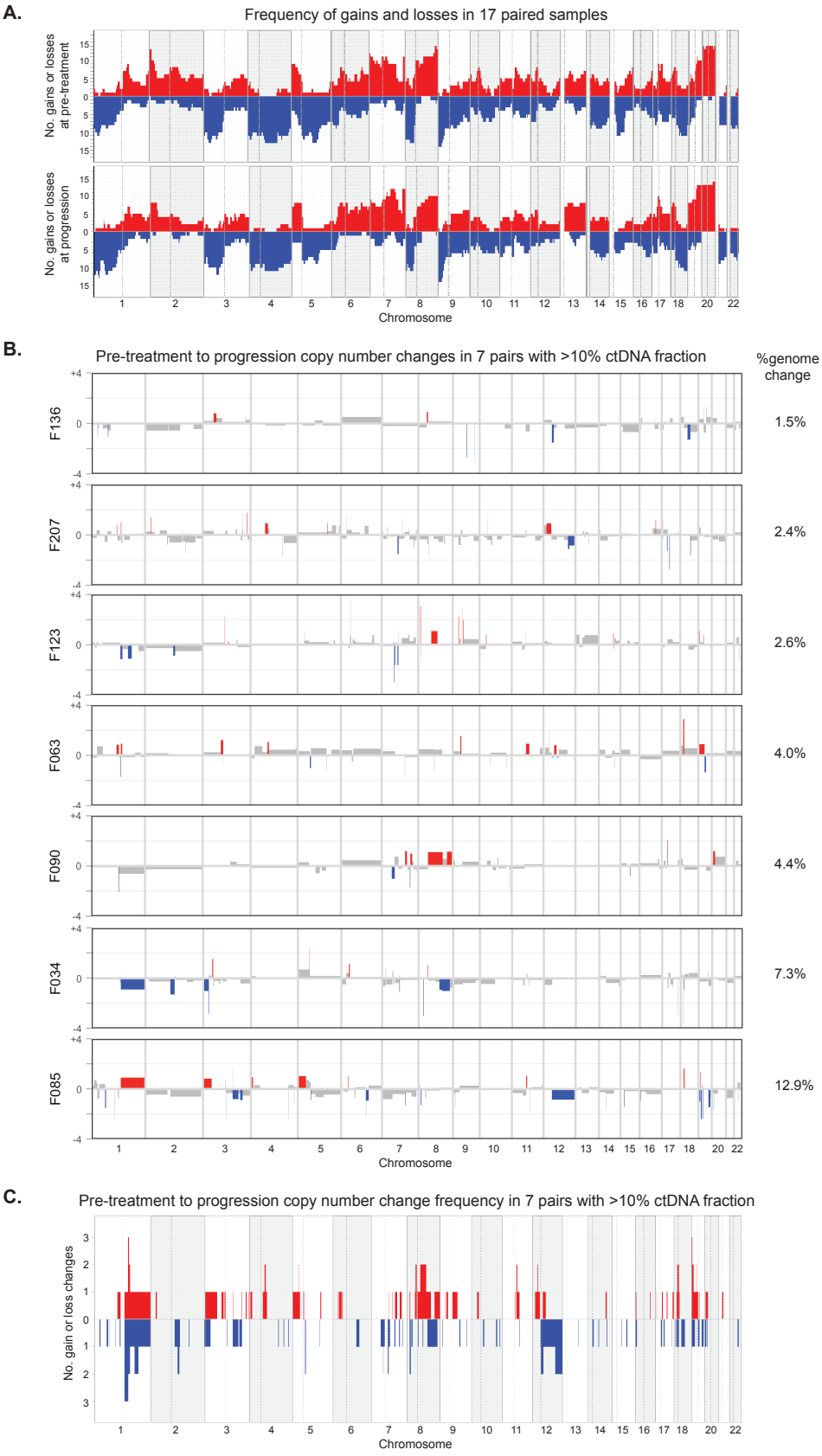
		N	Median ctDNA fraction (%)	p-value
All paired cases	Pre-treatment	20	15.18	0.1567
	Progression	20	8.72	
Initial radiological response followed by progression to chemotherapy: ‘primary responders’	Pre-treatment	12	17.00	0.0200
	Progression	12	7.59	
Stable disease or primary radiological progression to chemotherapy: ‘primary non-responders’	Pre-treatment	8	11.27	0.7984
	Progression	8	13.58	

lcWGS was applied to cfDNA collected at the time of radiological progression during or after first line treatment from 20 patients that had detectable ctDNA pre-treatment profiles and had a post-treatment sample available. Twelve of these had an initial radiological response with subsequent disease progression (primary responders). Eight showed stable disease or primary progression during chemotherapy (primary non-responders). In the primary responder group, the ichorCNA ctDNA fraction at progression was significantly lower than at pre-treatment (17% vs 7.6%; p=0.02; Table 3) whereas no significant change was observed in the primary non-responder group. Only three out of twenty samples taken at progression had a ctDNA content of zero (Cases 68, 99, 183), showing that ctDNA remains

detectable in the majority of tumours. The copy number profiles of the remaining 17 cases (Supplemental Figure 1) were assessed for changes over the course of chemotherapy treatment (Figure 4A). Using the 50 kb bin approach, all focal amplifications present before treatment were re-identified at progression (data not shown). No new focal amplifications were identified at progression.

In a second approach, we subtracted the pre-treatment absolute copy number (generated with ichorCNA) from the absolute copy number in the matched progression sample to assess which chromosomes changed through chemotherapy. To avoid artefacts from differences in tumour content, this pairwise comparison was only performed in seven cases where tumour content was similar and above 10% at both pre-treatment and progression. Only changes of the integer copy number value exceeding +/-0.8 were considered to enrich for new aneuploidies that had likely occurred in the majority of cells in the tumour and to avoid overinterpretation of changes in small subclones. The SCNA profiles were overall similar before treatment and at progression, but multiple individual segmental and arm-level changes were observed (Figure 4B). The fraction of the genome that changed (defined as the percent of the total genomic length that changed) was higher in responders (median: 5.65%, n=4) than in non-responders (median: 2.6%, n=3, Figure 4B) but this was not statistically significant. Individual cases showed new gains or losses in multiple chromosomes. However, most of the genomic regions that changed between pre-treatment and progression were only observed in a single case, and no large regions were changed in more than two cases (Figure 4C).

Figure 4 (**A**) Frequency plots showing the number of cases (n=17) that show segment gains (red) or losses (blue) at pre-treatment (top) and at progression (bottom). (**B**) For 7 pairs where both samples had >10% ctDNA fraction, comparative plots showing absolute copy number gains and losses at progression relative to pre-treatment, ordered by the extent of genomic change. The percent genomic change for each sample is indicated to the right of each plot. Red= gain, Blue= loss, Black= no change. A minimum of 0.8 copy number change was required to score a gain or a loss. (**C**) Frequency plot showing the number of cases (n=7) that show segment gains (red) or losses (blue) at progression relative to pre-treatment.


A.
Frequency of gains and losses in 17 paired samples
No. gains or losses at pre-treatment
No. gains or losses at progression
Chromosome
B.
Pre-treatment to progression copy number changes in 7 pairs with >10% ctDNA fraction
%genome change
F136
1.5%
F207
2.4%
F123
2.6%
F063
4.0%
F090
4.4%
F034
7.3%
F085
12.9%
Chromosome
C.
Pre-treatment to progression copy number change frequency in 7 pairs with >10% ctDNA fraction
No. gain or loss changes
Chromosome

Discussion

Through use of liquid biopsy we successfully analysed the SCNA profiles of 76.7% of 30 advanced OGAs. Serial analysis before and after first line chemotherapy was feasible in 85% of cases (17/20) that had detectable ctDNA prior to treatment. This demonstrates proof-of-concept that lcWGS of cfDNA can reveal genome wide SCNA profiles in the majority of patients with advanced OGA, for example to investigate novel prognostic or predictive biomarkers.

We identified several clinical characteristics that should support the selection of patients with a predictably higher cfDNA analysis success rates in future studies: the presence of liver metastases was associated with the highest ctDNA concentrations, whilst the ctDNA concentration was also significantly higher if the primary tumour was in situ. This may be the result of more aggressive tumours presenting with synchronous metastatic disease at baseline compared to those with metachronous metastases following resection. All seven cases with zero ctDNA pre-treatment either only had locally advanced disease or low metastatic burden. The use of such biomarkers to select OGA patients who are suitable for cfDNA sequencing may allow prioritizing these for liquid biopsy-based genotyping over sequencing of OGA tumour tissue, which has had moderate reported success rates due to technical challenges such as frequent low tumour content in endoscopic biopsies [26] [27]. With readily assessable clinical characteristics to identify suitable patients, cfDNA analysis could become the method of choice to assess amplifications for molecular stratification and particularly to longitudinally investigate SCNA evolution.

Neither pre-treatment total cfDNA concentration nor ctDNA tumour content correlated with survival, however, a low absolute plasma ctDNA concentration was significantly associated with better OS. A previous gastric cancer case series has described an association between baseline cfDNA and both relapse risk and adverse prognosis in the advanced disease setting [28], however larger studies are needed to validate the clinical utility of such metrics for optimisation of treatment and surveillance strategies [29].

High chromosomal instability (CIN) has been linked to poorer prognosis and drug sensitivity across a range of cancer types and to drug resistance *in vitro* [22] [30]. Application of several CIN metrics could not identify a correlation with chemotherapy response or survival in our cohort. This could indicate that CIN metrics may perform less well when generated from ctDNA, as this samples a summative copy number profile of the entire cancer population. Alternatively, these metrics may only weakly correlate with aggressiveness and treatment sensitivity and specific genetic aberrations, acquired as a consequence of CIN, may be more relevant in determining the response and outcome of

individual tumours. Although studies of larger cohorts may be able to reveal an association in the future, our results suggest that analysis of these CIN metrics in ctDNA is unlikely to be useful to predict individual patient outcomes in unselected patients undergoing first line chemotherapy.

For patients with evaluable ctDNA, multiple SCNAs could be identified in genes that are currently clinically relevant, or may become relevant to future practice. In samples with detectable ctDNA we identified all amplifications that had been found by previous targeted sequencing of matched FFPE tissue samples [26]. In seven cases, lcWGS found an additional nine focal amplifications in genes that had been analysed by targeted sequencing in tissue (*ERBB2, MET, KRAS, CDK4*) and where no amplification had been called. In three cases where tumour tissue sequencing failed, amplifications in *ERBB2, FGFR2, EGFR*, and *CCND1* were identified in ctDNA. Furthermore, lcWGS revealed multiple additional amplifications of potentially targetable driver genes such as *VEGFA*, highlighting the advantage of whole genome approaches over predetermined targeted sequencing gene sets.

Concurrent pre-treatment amplifications of *MAPK1*, *MET*, or *VEGFA* with *ERBB2* were seen in 3/6 HER2 positive cases. These may potentially influence variability of outcomes to HER2 targeted therapy, as amplifications of *MET* and *MAPK1* have previously been implicated in trastuzumab resistance [31] [32]. However the limited numbers in this cohort precluded meaningful survival analyses.

Comparison of pre-treatment SCNA profiles revealed gains of chromosomes 2q and 8p in cases that subsequently responded to treatment, and these gains were absent in non-responders. These need to be investigated in larger cohorts to assess their potential role as predictive biomarkers. The uniquely gained region on chromosome 8p harbours the DNA damage regulator *MCPH1,* which has been suggested to increase sensitivity to platinum chemotherapy [19]. This is therefore a candidate gene for further investigation. Identifying predictive biomarkers of chemotherapy response is an unmet need, but has been challenging; to date, the most extensive study of genetic predictors of therapy response using targeted sequencing of tumour tissue in advanced OGA failed to identify any biomarkers of response to platinum based chemotherapy [27].

Both ctDNA detection and lcWGS was possible from plasma samples taken at the timepoint of progression on first line chemotherapy, with 17/20 (85%) cases having detectable ctDNA. SCNA profiles were relatively stable between the pre-treatment and progression samples, but segmental and whole chromosomal arm changes were detected in seven cases where pair-wise comparison was quantifiable. As it is unlikely that multiple

subclones within a cancer would all gain or lose the same chromosomal regions, these copy number changes suggested that there had been major shifts in the clonal composition of the tumour cell populations with one or a few subclones becoming dominant whereas others had been lost. lcWGS may therefore be a useful technology for the investigation of resistance landscapes in larger cohorts. The lack of recurrent copy number change events at progression in this study may be a result of the small evaluable cohort, but equally it is feasible that chemotherapy resistance may be driven by point mutations. Use of a higher resolution technique that will allow the combined analysis of SCNA and mutations (such as whole exome cfDNA sequencing) may be warranted, with patient selection based on the presence of liver metastases to maximise successful sequencing rates and cost efficiency. Longitudinal cfDNA analysis has become a favoured method to interrogate resistance mechanisms during treatment, such as the tracking of known oncogenic *RAS* mutations in colorectal cancer [33]. Dynamic cfDNA testing should be equally applicable to monitor resistance to therapy in OGA.

The potential clinical application for this technique lies in the feasibility of biomarker stratification on the basis of lcWGS cfDNA sequencing, circumventing some of the limitations related to tumour heterogeneity in OGA [13]. Furthermore, sequential lcWGS of cfDNA is a low cost method for continuing to investigate genetic changes associated with chemotherapy response in larger series or for early detection of resistance mechanisms to novel agents in clinical trials. Preliminary proof of concept for the use of longitudinal cfDNA analysis to predict response and resistance to HER2-targeting treatment has already been described [34]: *ERBB2* copy number alterations detected by targeted sequencing were found to be associated with both innate and acquired trastuzumab resistance. Additionally, mutations in genes including *PIK3CA*, *ERBB2* and *ERBB4* were also associated with resistance, highlighting the benefit of combined mutation identification and SCNA analysis in interrogating drug resistance mechanisms. Detection of relevant gene amplifications in cfDNA has been already shown to be clinically important for patient selection and therapeutic targeting of FGFR in gastric cancer [13]. However, the plasma contains multiple components in addition to cfDNA that could also be utilised to realise the full potential of the liquid biopsy. Promising techniques under investigation in OGA include the enumeration and characterization of circulating tumour cells (CTCs), which has been associated with both prognosis [35] and treatment response [36]. In prostate cancer, mRNA extracted from CTCs has been used to identify splice variants of the androgen receptor that are prognostic for taxane therapy [37]. Furthermore, CTCs from small cell lung cancer have been successfully cultured ex vivo in order to screen for targeted therapy sensitivity and relevant biomarkers

[38, 39]. As an alternative to CTCs and cell free nucleic acids, exosomes may also provide a means for tumour profiling, including in OGA [40].

As novel targeted and immune-modulating therapies are introduced into clinical management of OGA, there will be a need for stratification of patients in order to guide personalised treatment. The use of genome-wide analysis to interrogate key driver events and genomic evolution over time will be important in refining the effective biomarker stratification of such treatments moving forwards. It is possible that a combination of lcWGS cfDNA sequencing with CTC or exosome analyses will facilitate maximal clinical utility to be gained from liquid biopsy approaches in order to guide treatment decisions. Ultimately this may support precision medicine in both trial and routine clinical practice settings by avoiding the cost, delay and clinical complications of repeated invasive biopsy procedures.

Methods

Trial design and sample collection

The FOrMAT (Feasibility of a Molecular Characterisation Approach to Treatment, Chief Investigator: N Starling ClinicalTrials.gov NCT02112357) study enrolled patients with advanced gastrointestinal malignancies treated at the Royal Marsden from February 2014 to November 2015 [26]. The trial was approved by the UK National Ethics Committee (approval number: 13/LO/1274RM) and all patients provided written informed consent. As part of the tissue collection component of the trial, blood samples were obtained at trial entry and at the timepoint of response assessment CT scans during treatment. The trial recruited 71 advanced OGA cancer patients in total. The clinical trial database was interrogated to identify 30 patients with a diagnosis of locally advanced inoperable or metastatic OGA who had undergone pre-treatment research blood sampling prior to commencement of treatment, and whom had sequential bloods spanning at least the full course of comparable first-line systemic chemotherapy, consisting of a platinum/fluoropyrimidine doublet in all cases, plus or minus anthracyline or, in the case of *ERBB2* positive tumours, trastuzumab. cfDNA was extracted from plasma samples taken at a baseline pre-treatment timepoint for all patients. To assess the evolution of SCNA through treatment, lcWGS was additionally performed on cfDNA collected at the time of radiological progression during or after first line platinum and 5FU based combination chemotherapy from 20 patients that had detectable ctDNA pre-treatment profiles and had a post-treatment sample available.

cfDNA extraction and quantification

Plasma was separated within 2 hours of blood draw and frozen at -80C. The QIAamp Circulating Nucleic Acid Kit (Qiagen) was used to isolate cfDNA from 3-4 mL plasma according to manufacturer's instructions. cfDNA within a size range of 100 to 700bp was quantified using a Bioanalyzer High Sensitivity chip (Agilent), encompassing the predominant 3 cfDNA fragment peaks [41].

lcWGS

For the majority of cases 10ng of input DNA was used for sequencing, although in some cases with limited yield, 5ng were used [42]. Libraries were prepared using the NEBNext Ultra DNA Library Prep kit (NEB), pooled and sequenced on an Illumina HiSeq2500 in Rapid mode single read 100bp.

SCNA analysis

Sequencing reads were aligned to the human reference genome (hg19) using Bowtie (v1.2.9) [43] and resultant bam files were deduplicated using Picard *MarkDuplicates* (http://picard.sourceforge.net; v.2.1.0). Reads were subsequently assigned to non-overlapping 500 kb bins and normalized to correct for GC-content and mappability bias using the HMMcopy suite (http://compbio.bccrc.ca/software/hmmcopy/) [44]. IchorCNA [17] was used to quantify tumour fraction in cfDNA from lcWGS without prior knowledge of SSNV or SCNAs present in the primary tumour sample. IchorCNA segmented data was normalised using the best-fit tumour content and ploidy solution in order to compare samples. To compare multiple samples, data was uniformally segmented using interpolate.pcf, part of the copynumber package in R (http://bioconductor.org/packages/copynumber/) [45]. Cohort frequency plots were generated using the copynumber plotFreq function. Seg files were viewed as a heat map using the Integrated Genome Viewer software (Broad Institute; v.2.3.97), allowing comparison of genomic SCNA profiles across multiple samples with the ability to zoom in to areas of interest in order to investigate genes located within this genomic region [46]. Focal SCNAs were identified by assigning mapped reads to 50kb bins using the method described by Baslan [25]. SCNAs were assessed in IGV by two independent observers and recorded for all patients.

Survival analyses by pre-treatment circulating DNA metrics

Tertile survival analyses were undertaken according to three circulating DNA metrics: (1) total cfDNA concentration extracted from plasma, (2) ctDNA content estimated by ichorCNA and (3) absolute ctDNA concentration in the plasma, calculated by multiplying the total cfDNA concentration with the ichorCNA ctDNA content. In each case the 30 samples

were classified into ‘low’, ‘medium’ and ‘high’ tertiles for each metric and overall survival trend was analysed using log-rank method.

Data availability

Sequence reads have been deposited in the European Genome Phenome Archive (ID: submission ongoing – will be updated as soon as ID assigned).

Conclusions

SCNA profiles were successfully analysed through the use of lcWGS applied to cfDNA extracted from pre-treatment baseline plasma samples in 23/30 (76.7%) cases. The presence of liver metastases, primary tumour in situ and oesophageal or junctional primary tumour site were associated with higher pre-treatment ctDNA content, and a lower baseline ctDNA concentration was associated with subsequent improved overall survival. Concordance was noted with prior targeted tumour sequencing results, and additionally lcWGS revealed additional amplifications of potentially targetable driver genes, highlighting the advantage of whole genome approaches over predetermined targeted sequencing gene sets. ctDNA detection and lcWGS was possible from plasma samples taken at the timepoint of progression on first line chemotherapy, with SCNA profiles successfully analysed in 17/20 (85%) cases. Although SCNA profiles were relatively stable between pre-treatment and progression, segmental and whole chromosomal arm changes were detected in seven cases where pair-wise comparison was quantifiable. Such shifts in the clonal composition of tumour cell populations during chemotherapy warrant further investigation as a possible dynamic means of investigating resistance landscapes in OGA.

Acknowledgements

We would like to thank all patients participating in the FOrMAT clinical trial and the clinical research team members at the Royal Marsden Hospital who supported the sample collection. The study was supported by the National Institute for Health Research Biomedical Research Centre for Cancer at the ICR/RMH. MG, LB, AW, and BG were supported by CRUK, a charitable donation from Tim Morgan, Cancer Genetics UK and the Constance Travis Trust. The ICR Centre for Evolution and Cancer was supported by a Wellcome Trust Strategic Grant (105104/Z/14/Z).

Author Contributions

Conceptualization, MD, LJB and MG; Formal analysis, MD, LJB, AW and MG; Funding acquisition, NS and MG; Investigation, MD, LJB, CC, SM and BG; Methodology, LJB and IA; Project administration, MG; Resources, SYM, IR, RB, NM, SR, DW, IC, DC and NS; Writing – original draft, MD, LJB and MG; Writing – review & editing, NS.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Bray F, Ferlay J, Soerjomataram I et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA. Cancer J. Clin. 2018. doi:10.3322/caac.21492.

2. Wagner AD, Unverzagt S, Grothe W et al. Chemotherapy for advanced gastric cancer. Cochrane database Syst. Rev. 2010; 3(3):CD004064.

3. Bass AJ, Thorsson V, Shmulevich I et al. Comprehensive molecular characterization of gastric adenocarcinoma. Nature 2014; 513(7517):202–9.

4. Kim J, Bowlby R, Mungall AJ et al. Integrated genomic characterization of oesophageal carcinoma. Nature 2017; 541(7636):169–175.

5. Liang L, Fang J-Y, Xu J. Gastric cancer and gene copy number variation: emerging cancer drivers for targeted therapy. Oncogene 2016; 35(12):1475–1482.

6. Labots M, Buffart TE, Haan JC et al. High-level copy number gains of established and potential drug target genes in gastric cancer as a lead for treatment development and selection. Cell. Oncol. 2014; 37(1):41–52.

7. Zhang Y. Epidemiology of esophageal cancer. World J. Gastroenterol. 2013; 19(34):5598–606.

8. Zhou B, Ho SS, Zhang X et al. Whole-genome sequencing analysis of CNV using low-coverage and paired-end strategies is efficient and outperforms array-based CNV analysis. J. Med. Genet. 2018; 55(11):735 LP – 743.

9. Heitzer E, Auer M, Hoffmann EM et al. Establishment of tumor-specific copy number alterations from plasma DNA of patients with cancer. Int. J. cancer 2013; 133(2):346–56.

10. Alsina M, Gullo I, Carneiro F. Intratumoral heterogeneity in gastric cancer: a new challenge to face. Ann. Oncol. 2017; 28(5):912–913.

11. Murugaesu N, Wilson GA, Birkbak NJ et al. Tracking the genomic evolution of esophageal adenocarcinoma through neoadjuvant chemotherapy. Cancer Discov. 2015; 5(8):821–832.

12. Lee HE, Park KU, Yoo SB et al. Clinical significance of intratumoral HER2 heterogeneity in gastric cancer. Eur. J. Cancer 2013; 49(6):1448–57.

13. Pearson A, Smyth E, Babina IS et al. High-Level Clonal FGFR Amplification and Response to FGFR Inhibition in a Translational Clinical Trial. Cancer Discov. 2016; 6(8):838–851.

14. Petty RD, Dahle-Smith A, Stevenson DAJ et al. Gefitinib and EGFR Gene Copy Number Aberrations in Esophageal Cancer. J. Clin. Oncol. 2017; 35(20):2279–2287.

15. Gao J, Wang H, Zang W et al. Circulating tumor DNA functions as an alternative for tissue to overcome tumor heterogeneity in advanced gastric cancer. Cancer Sci. 2017; 108(9):1881–1887.

16. Pectasides E, Stachler MD, Derks S et al. Genomic Heterogeneity as a Barrier to Precision Medicine in Gastroesophageal Adenocarcinoma. Cancer Discov. 2018; 8(1):37 LP – 48.

17. Adalsteinsson VA, Ha G, Freeman SS et al. Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. Nat. Commun. 2017; 8(1):1324.

18. Chaplet M, Rai R, Jackson-Bernitsas D et al. BRIT1/MCPH1: A guardian of genome and an enemy of tumors. Cell Cycle 2006. doi:10.4161/cc.5.22.3471.

19. Wei Q.; Wang, X.; An, X.; Han, Q.; Meng, L.; Cao, W. L. Z. Effects of MCPH1 silencing on proliferation, apoptosis, and chemo-sensitivity of non-small cell lung cancer cells. Int. J. Clin. Exp. Med. 2018.

20. Dulak AM, Schumacher SE, Van Lieshout J et al. Gastrointestinal adenocarcinomas of the esophagus, stomach, and colon exhibit distinct patterns of genome instability

and oncogenesis. Cancer Res. 2012. doi:10.1158/0008-5472.CAN-11-3893.

21. Dewhurst SM, McGranahan N, Burrell RA et al. Tolerance of whole- genome doubling propagates chromosomal instability and accelerates cancer genome evolution. Cancer Discov. 2014. doi:10.1158/2159-8290.CD-13-0285.

22. Vargas-Rondón N, Villegas EV, Rondón-Lagos M. The Role of Chromosomal Instability in Cancer and Therapeutic Responses. Cancers 2018. doi:10.3390/cancers10010004.

23. Burrell RA, Mcgranahan N, Bartek J, Swanton C. The causes and consequences of genetic. Nature 2013; 501:338–345.

24. Endesfelder D, Burrell RA, Kanu N et al. Chromosomal instability selects gene copy-number variants encoding core regulators of proliferation in ER+ Breast cancer. Cancer Res. 2014; 74(17):4853–4863.

25. Baslan T, Kendall J, Rodgers L et al. Genome-wide copy number analysis of single cells. Nat. Protoc. 2012; 7(6):1024–1041.

26. Moorcraft SY, Gonzalez de Castro D, Cunningham D et al. Investigating the feasibility of tumour molecular profiling in gastrointestinal malignancies in routine clinical practice. Ann. Oncol. 2018; 29(1):230–236.

27. Janjigian YY, Sanchez-Vega F, Jonsson P et al. Genetic Predictors of Response to Systemic Therapy in Esophagogastric Cancer. Cancer Discov. 2018; 8(1):49–58.

28. Fang W-L, Lan Y-T, Huang K-H et al. Clinical significance of circulating plasma DNA in gastric cancer. Int. J. Cancer 2016; 138(12):2974–2983.

29. Saluja H, Karapetis CS, Pedersen SK et al. The Use of Circulating Tumor DNA for Prognosis of Gastrointestinal Cancers. Front. Oncol. 2018; 8:275.

30. Lee AJX, Endesfelder D, Rowan AJ et al. Chromosomal instability confers intrinsic multidrug resistance. Cancer Res. 2011; 71(5):1858–1870.

31. Pietrantonio F, Fuca G, Morano F et al. Biomarkers of primary resistance to trastuzumab in HER2-positive metastatic gastric cancer patients: The AMNESIA case-control study. Clin. Cancer Res. 2018; 24(5):1082–1089.

32. Menyhart O, Santarpia L, Gyorffy B. A Comprehensive Outline of Trastuzumab Resistance Biomarkers in HER2 Overexpressing Breast Cancer. Curr. Cancer Drug

606 Targets 2015; 15(8):665–683.

607 33. Berger AW, Schwerdel D, Welz H et al. Treatment monitoring in metastatic colorectal
608 cancer patients by quantification and KRAS genotyping of circulating cell-free DNA.
609 PLoS One 2017; 12(3):e0174308.

610 34. Wang D-S, Liu Z-X, Lu Y-X et al. Liquid biopsies to track trastuzumab resistance in
611 metastatic HER2-positive gastric cancer. Gut 2018:gutjnl-2018-316522.

612 35. Zheng X, Fan L, Zhou P et al. Detection of Circulating Tumor Cells and Circulating
613 Tumor Microemboli in Gastric Cancer. Transl. Oncol. 2017; 10(3):431–441.

614 36. Mishima Y, Matsusaka S, Chin K et al. Detection of HER2 Amplification in Circulating
615 Tumor Cells of HER2-Negative Gastric Cancer Patients. Target. Oncol. 2017;
616 12(3):341–351.

617 37. Tagawa ST, Antonarakis ES, Gjyrezi A et al. Expression of AR-V7 and ARV 567Es in
618 circulating tumor cells correlates with outcomes to taxane therapy in men with
619 metastatic prostate cancer treated in taxynergy. Clin. Cancer Res. 2019.
620 doi:10.1158/1078-0432.CCR-18-0320.

621 38. Lallo A, Frese KK, Morrow CJ et al. The combination of the PARP inhibitor olaparib
622 and the WEE1 Inhibitor AZD1775 as a new therapeutic option for small cell lung
623 cancer. Clin. Cancer Res. 2018. doi:10.1158/1078-0432.CCR-17-2805.

624 39. Lallo A, Gulati S, Schenk MW et al. Ex vivo culture of cells derived from circulating
625 tumour cell xenograft to support small cell lung cancer research and experimental
626 therapeutics. Br. J. Pharmacol. 2019. doi:10.1111/bph.14542.

627 40. Fu M, Gu J, Jiang P et al. Exosomes in gastric cancer: Roles, mechanisms, and
628 applications. Mol. Cancer 2019. doi:10.1186/s12943-019-1001-7.

629 41. Mansukhani S, Barber LJ, Kleftogiannis D et al. Ultra-Sensitive mutation detection
630 and genome-wide DNA copy number reconstruction by error- corrected circulating
631 tumor DNA sequencing. Clin. Chem. 2018; 64(11):1626–1635.

632 42. Heitzer E, Ulz P, Belic J et al. Tumor-associated copy number changes in the
633 circulation of patients with prostate cancer identified through whole-genome
634 sequencing. Genome Med. 2013; 5(4):30.

635 43. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient
636 alignment of short DNA sequences to the human genome. Genome Biol. 2009.

637 doi:10.1186/gb-2009-10-3-r25.

638 44. Ha G, Roth A, Lai D et al. Integrative analysis of genome-wide loss of heterozygosity
639 and monoallelic expression at nucleotide resolution reveals disrupted pathways in
640 triple-negative breast cancer. Genome Res. 2012; 22(10):1995–2007.

641 45. Nilsen G, Liestøl K, Loo P Van et al. Copynumber: Efficient algorithms for single- and
642 multi-track copy number segmentation. BMC Genomics 2012. doi:10.1186/1471-
643 2164-13-591.

644 46. Robinson JT, Thorvaldsdóttir H, Winckler W et al. Integrative genomics viewer. Nat.
645 Biotechnol. 2011; 29:24.

646

647

648

Supplemental Table: Genes in frequently gained region of chromosome 8p in responders

CSMD1
LOC100287015
MCPH1
ANGPT2
CLDN23
MFHAS1
ERI1
MIR4660
PPP1R3B
LOC157273
TNKS
MIR597
LINC00599
MIR124-1
MSRA
PRSS55
RP1L1

Supplemental Figure 1 Integer copy number profiles for the 17 paired non-zero ctDNA cases at progression. ichorCNA ctDNA fraction is indicated for each sample.