


RESEARCH

Open Access



# Molecular evolutionary process of advanced gastric cancer during sequential chemotherapy detected by circulating tumor DNA

Wenqi Xi<sup>1†</sup>, Chenfei Zhou<sup>1,2†</sup>, Fei Xu<sup>4</sup>, Debin Sun<sup>4</sup>, Shengzhou Wang<sup>4</sup>, Yawei Chen<sup>4</sup>, Jun Ji<sup>3</sup>, Tao Ma<sup>5</sup>, Junwei Wu<sup>1,2</sup>, Chengfang Shangguan<sup>1</sup>, Zhenggang Zhu<sup>1,3</sup> and Jun Zhang<sup>1,2,6\*</sup> 

## Abstract

**Background:** Efficacy of conventional sequential chemotherapy paradigm for advanced gastric cancer (AGC) patients has largely plateaued. Dynamic molecular changes during and after sequential chemotherapy have not been fully delineated. We aimed to profile the molecular evolutionary process of AGC patients during sequential chemotherapy by next generation sequencing (NGS) of plasma circulating tumor DNA (ctDNA).

**Methods:** A total of 30 chemo-naïve patients who were diagnosed with unresectable advanced or metastatic stomach adenocarcinoma were enrolled. All patients received sequential chemotherapy regimens following the clinical guideline. One hundred and eight serial peripheral blood samples were collected at baseline, radiographical assessment and disease progression. Plasma ctDNA was isolated and a customized NGS panel was used to detect the genomic features of ctDNA including single nucleotide variants (SNVs) and gene-level copy number variations (CNVs). KEGG pathway enrichment analysis was performed.

**Results:** Platinum-based combination chemotherapy was administrated as first-line regimen. Objective response rate was 50% (15/30). Patients with higher baseline values of copy number instability (CNI), CNVs and variant allele frequency (VAF) were more sensitive to platinum-based first-line regimens. Tumor mutation burden (TMB), CNI and CNV burden at partial response and stable disease were significantly lower than those at baseline, where at progressive disease they recovered to baseline levels. Dynamic change of TMB ( $\Delta$ TMB) was correlated with progression-free survival of first-line treatment. Fluctuating changes of SNVs and gene-level CNVs could be observed during sequential chemotherapy. Under the pressure of conventional chemotherapy, the number of novel gene-level CNVs were found to be higher than that of novel SNVs. Such novel molecular alterations could be enriched into multiple common oncologic signaling pathways, including EGFR tyrosine kinase inhibitor resistance and platinum drug resistance pathways, where their distributions were found to be highly heterogenous among patients. The impact of subsequent regimens, including paclitaxel-based and irinotecan-based regimens, on the molecular changes driven by first-line therapy was subtle.

<sup>†</sup>Wenqi Xi and Chenfei Zhou have contributed equally to this manuscript and should be considered as co-first authors

\*Correspondence: junzhang10977@sjtu.edu.cn

<sup>1</sup> Department of Oncology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, No. 197 Ruijin er Road, Shanghai 200025, China  
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Conclusion:** Baseline and dynamic changes of genomic features of ctDNA could be biomarkers for predicting response of platinum-based first-line chemotherapy in AGC patients. After treatment with standard chemotherapy regimens, convergent oncologic pathway enrichment was identified, which is yet characterized by inter-patient heterogeneous gene-level CNVs.

**Keywords:** Gastric cancer, Sequential chemotherapy, Dynamic ctDNA analysis, Copy number variations, Tumor evolution

## Introduction

Gastric cancer is one of the leading causes of cancer death in China with 5-year overall survival (OS) that is barely over 30% [1, 2]. The backbone therapeutic for advanced gastric cancer (AGC) patients remains to be multi-drug combination chemotherapy, which has shown efficacy that has largely plateaued [3]. The median progression-free survival (PFS) of AGC patients is about 6–8 months in first-line setting. Despite progress made in understanding mechanisms that explain chemotherapy resistance of AGC using bioinformatics and multi-omics analysis technologies, biomarkers deemed to be able to identify patients who are more likely to respond to combination chemotherapy are surrounded by controversy [4–6].

Options available to AGC patients as subsequent regimens after first-line treatment are limited [7–9], where the efficacy of current treatment patterns is unsatisfactory. Under the pressure of chemotherapy, tumor cells may evolve to adapt to chemotherapy and acquire a resistant phenotype. Fitness trade-offs occur during this process, which could induce vulnerability to some other drugs. Thus, understanding the molecular evolution in place during the treatment may help to optimize the subsequent treatment. Re-biopsy for late-stage patients is normally difficult, and high inter- and intra-tumoral heterogeneity of gastric cancer suggests that genomic analysis based on tissue biopsy may hardly provide results that are representative of the molecular alterations of the tumor.

Analysis of circulating tumor DNA (ctDNA) using next generation sequencing (NGS) offers a mini-invasive method to avoid the intra-tumoral heterogeneity and to dynamically monitor the molecular alterations in place [10–12]. The correlation between the genomic alterations of ctDNA and the clinical efficacy of the systemic treatment of gastric cancer has already been recognized in the literature [13]. Changes of quantitative data including ctDNA concentration, mutant allele fraction, molecular tumor burden index which have been reported by several research groups can provide indications of cancer recurrence, overall survival, response to systemic therapy of gastrointestinal cancer patients [12, 14, 15]. Somatic mutations of specific genes like *TGFBR2*, *RHOA*,

*ERBB2* in ctDNA were found as potential biomarkers of the response to immunotherapy or HER2-targeted therapies [16–18]. However, most of these studies focused on genomic alterations of ctDNA at a specific timepoint during treatment. The dynamic molecular changes and their potential clinical meaning during the sequential chemotherapy in AGC have not been fully delineated.

In the present study, peripheral blood samples were serially and prospectively collected from AGC patients who undergone treatments by sequential chemotherapy regimens. Plasma ctDNA were analyzed utilizing a customized NGS panel. Our aims were to identify the correlations between the genomic features of ctDNA and the patients' outcomes and to profile the molecular evolutionary process of AGC patients during the sequential chemotherapy to further explore the potential therapeutic options available to these patients.

## Results

### Clinical characteristics of enrolled patients

A total of 30 patients (20 males and 10 females) were enrolled in the study. The patients' clinical characteristics are listed in Table 1. Seventeen cases had hematogenous metastases, including lung, liver, bone, and adrenal gland. Two chemotherapy regimens were administrated as first-line settings, which included platinum-based triplet regimen (Pt-3d; 14 cases) and platinum-based doublet regimen (Pt-2d; 16 cases) [19]. The details of the treatment results are listed in Additional file 8: Table S1. The best objective response rate of the first-line treatment was 50% (15/30). At the end of the follow-up (January 1st, 2021), four patients were still alive. The median overall survival (OS) of these patients was 9.5 months (range: 2.8–27.6) and the median progression-free survival (PFS) of the first-line setting was 4.8 months (range: 1.5–41.7). A total of 108 peripheral blood samples were eligible for ctDNA analysis, and the median sampling times of the patients was 3 (range 2–10).

### Baseline genomic features of ctDNA correlated with treatment response of first-line regimens

Patients were stratified into R (responsive) and NR (non-responsive) groups by their best objective response to first-line treatment. The clinical characteristics of the

**Table 1** Clinical characteristics of 30 AGC patients

Clinical characteristics	Case (n)	Percentage
Gender		
Male	20	66.7
Female	10	33.3
Age		
Median	64.0	
Range	37–77	
Tumor site		
Fundus	6	20.0
Body	10	33.3
Antrum	14	46.7
Hematogenous metastasis		
Yes	17	56.7
No	13	43.3
Regimens		
Pt-3d	14	46.7
FLOT	11	–
DOS	3	–
Pt-2d	16	53.3
SOX	12	–
CAPOX	3	–
FOLFOX	1	–
PTX	5	16.7
Paclitaxel	2	–
PX	2	–
Nab-paclitaxel	1	–
CPT-11	6	20.0
IRIR	4	–
Irinotecan	1	–
FOLFIRI	1	–
Treatment pattern		
Pt-3d to CPT-11	6	20.0
Pt-2d to PTX	5	16.7
ORR of 1st-line		
PR	15	50.0
SD	10	33.3
PD	5	16.7

AGC advanced gastric cancer

two groups were not significantly different (Additional file 8: Table S2). At baseline, compared to NR group, R group showed significantly higher copy number instability (CNI) value ( $4070.3 \pm 1648.7$  vs  $2406.2 \pm 1107.3$ ;  $P=0.00058$ ) and copy number variation (CNV) burden ( $45 \pm 37$  vs  $13 \pm 21$ ,  $P=0.0036$ ), whereas values of ctDNA content fraction (CCF) and tumor mutation burden (TMB) were not found to be significantly different (Fig. 1A). Mean variant allele frequency (meanVAF) and maximum variant allele frequency (maxVAF) were

also higher in R group (Additional file 1: Figure S1A). The best cut-off value of CNI to stratify R and NR groups was 2339.33 (Additional file 1: Figure S1B). Patients with CNI-high showed similar PFS ( $P=0.78$ ) and OS ( $P=0.89$ ) to those of patients with CNI-low (Additional file 1: Figure S1C).

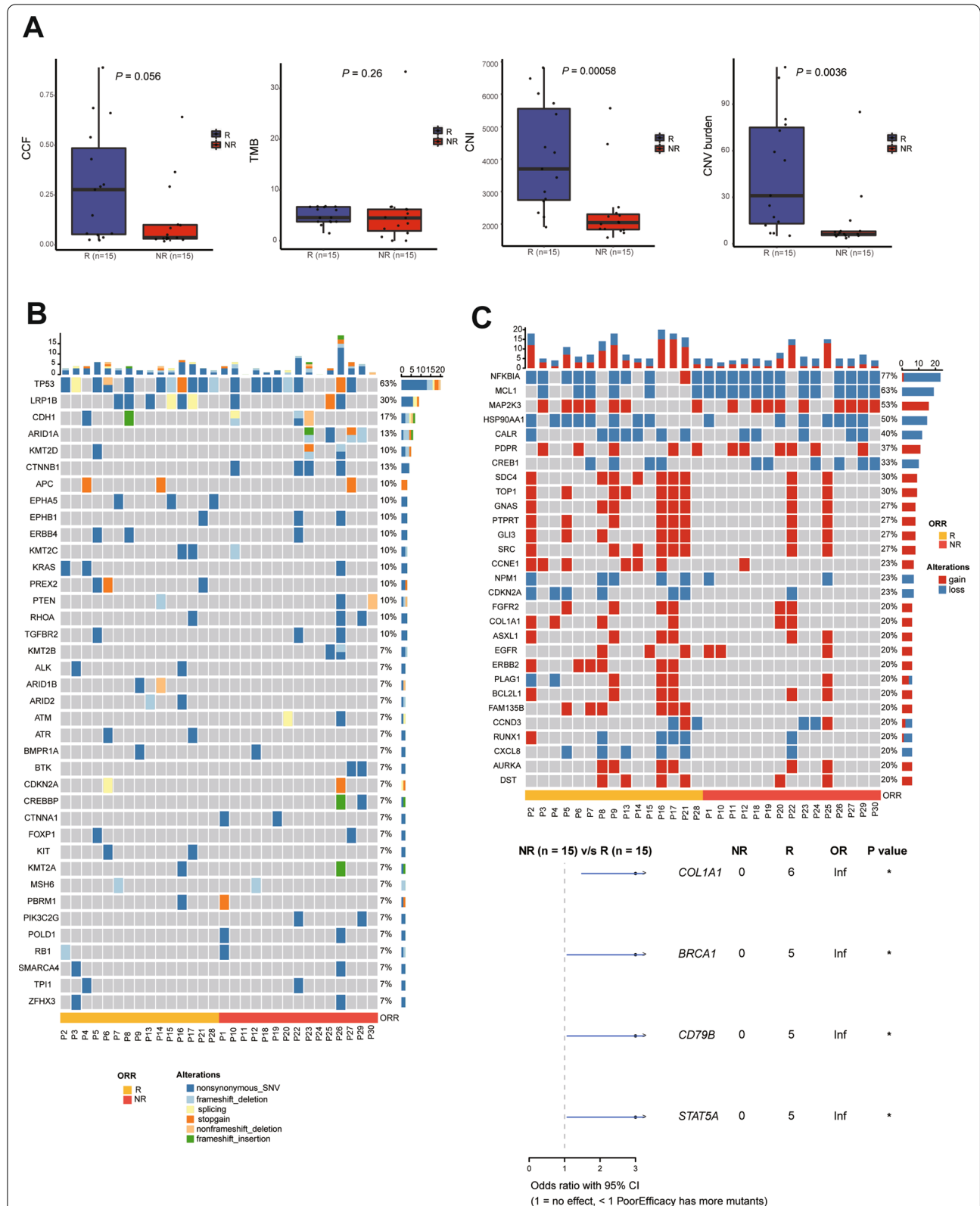
At baseline, top 10 somatic single nucleotide variants (SNVs) were observed in *TP53* (63%), *LRP1B* (30%), *CDH1* (17%), *ARID1A* (13%), *CTNNA1* (13%), *KMT2D* (10%), *APC* (10%), *EPHA5* (10%), *EPHB1* (10%) and *ERBB4* (10%). The distributions of SNVs in R and NR groups were similar (Fig. 1B). The top 10 gene-level CNVs were observed in *NFKB1A* (77%), *MCL1* (63%), *MAP2K3* (53%), *HSP90AA1* (50%), *CALR* (40%), *PDPK1* (37%), *CREB1* (33%), *SDC4* (30%), *TOP1* (30%) and *GNAS* (27%). CNVs in R group were more frequent than those in NR group (Fig. 1C). Copy number gain in *COL1A1*, *BRCA1*, *CD79B* and *STAT5A* was only found in R group and was significantly correlated with treatment response of first-line therapy (Fig. 1C). The summarized molecular information on SNV and CNV was shown in Additional file 9: Table S3 and Additional file 10: Table S4, respectively.

#### Dynamic change of genomic alterations in ctDNA reflected clinical efficacy during chemotherapy

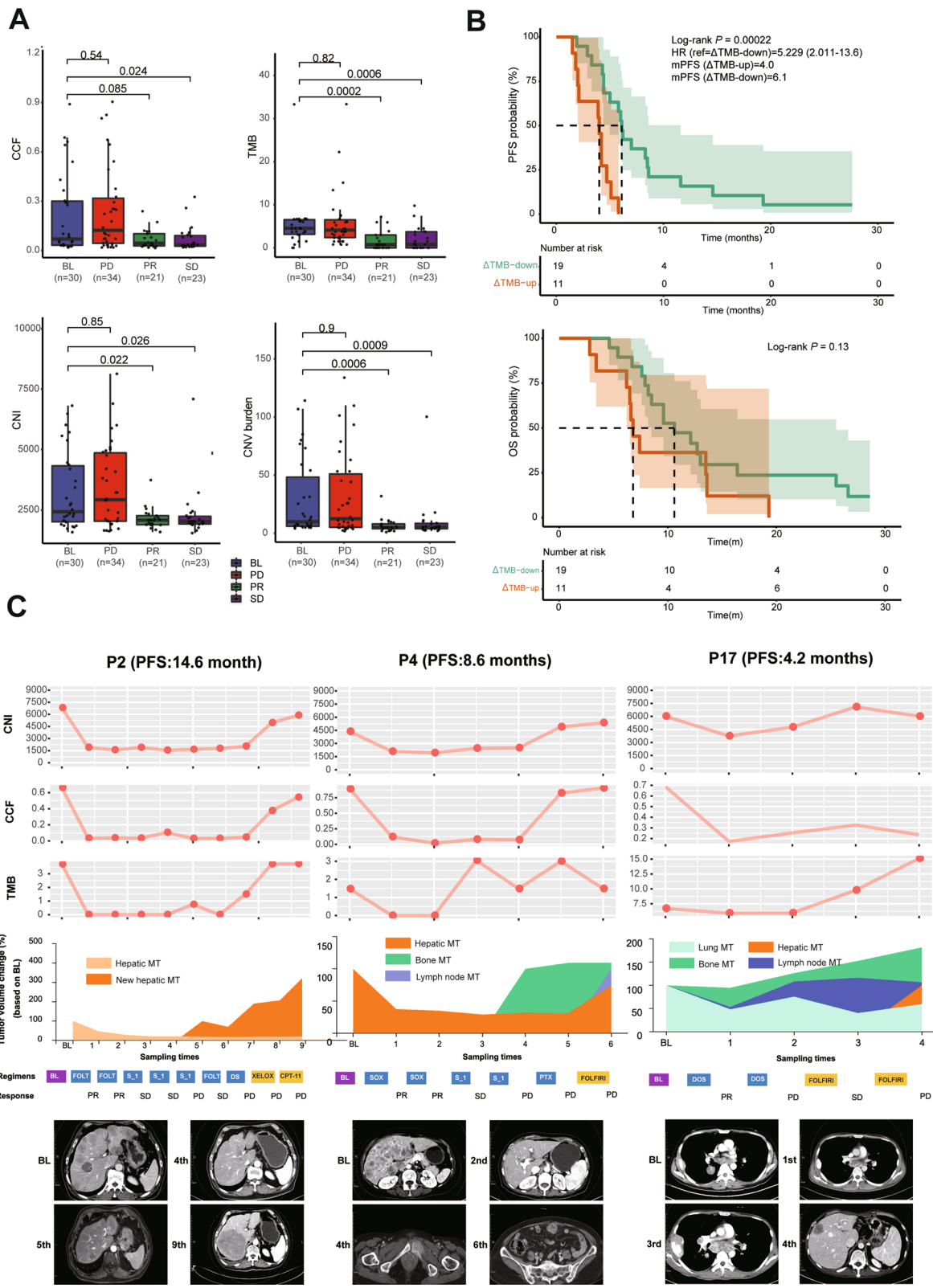
Values of TMB, CNI and CNV burden at partial response (PR) and stable disease (SD) were significantly lower than those samples at baseline (BL), whereas at progressive disease (PD) they were similar to BL values. The value of CCF at SD was also lower than that at BL, whereas no significant difference was observed between PR and BL (Fig. 2A). Values of meanVAF and maxVAF were not significantly different during the treatment (Additional file 2: Figure S2A).

The correlation between early change of genomic features (i.e., delta values between 2nd and 1st sampling points) and the outcome of patients was analyzed. The cut-off value of  $\Delta$ TMB was defined as  $-1$ , where 19 patients were defined as  $\Delta$ TMB-down ( $\Delta$ TMB  $< -1$ ) and 11 patients were defined as  $\Delta$ TMB-up ( $\Delta$ TMB  $> -1$ ). PFS of  $\Delta$ TMB-down patients was longer than that of  $\Delta$ TMB-up patients (6.1 months vs. 4.0 months;  $P < 0.001$ ), where OS was similar between two groups ( $P=0.13$ ; Fig. 2B). For other genomic features, no correlation with PFS was identified (Additional file 3: Figure S3).

The dynamic changes of genomic features in ctDNA and tumor volumes are described in three typical cases (Fig. 2C). The values of CNI, CCF and TMB all decreased after the initial response to the first-line treatment and were kept at low level during the disease control. After disease progression, CNI, CCF and TMB continued to increase during the subsequent treatment.



**Fig.1** Baseline genomic features of ctDNA correlated with treatment response of first-line regimens. **A:** values of CCF, TMB, CNI and CNV burden in R and NR groups. **B:** landscape of baseline SNVs in all patients. **C:** landscape of baseline CNVs in all patients and four genes correlated with R group



**Fig.2** Dynamic change of genomic features in ctDNA correlated with clinical efficacy during chemotherapy. **A:** values of CCF, TMB, CNI and CNV burden at baseline (BL), progressive disease (PD) and during treatment (PR, SD). **B:** correlation between ΔTMB of ctDNA and patients' outcomes of first-line therapy. **C:** dynamic changes of genomic features in ctDNA and tumor volumes in three typical cases



### Fluctuating changes of genomic alterations during sequential chemotherapy of AGC

Fluctuating changes could be observed once comparisons were made in SNVs and CNVs landscape patterns of 30 patients. Some SNVs (Additional file 2: Figure S2B) and CNVs (Additional file 2: Figure S2C) were undetectable at PR and SD, which were recovered at PD. *TP53* and *LRP1B* were top two genes with somatic mutations. The percentages of detectable *TP53* and *LRP1B* mutations were lower at PR compared to BL, where they increased at PD (Fig. 3A). For an individual patient, the mutation types and sites were the same during the whole course of the treatment (Fig. 3B). However, the sites of point mutations of these two genes were highly heterogeneous among the patients (Fig. 3C).

### Enrichment of novel genomic alterations of ctDNA during treatment as revealed by KEGG analysis

Novel SNVs occurred in 130 genes during the whole course of the treatment. Only 26 genes were found in more than 2 patients. The most frequently novel SNVs were observed in *KMT2C* (5%), *MED12* (5%), *PPM1D* (5%), *CHEK2* (4%), *FAT4* (4%), *KMT2D* (4%), *MMP8* (4%) and *PRKCI* (4%). Novel CNVs occurred in 191 genes, and 109 of them were observed in more than 2 patients. The most frequently novel CNVs were observed in *MCL* (28%), *CREB1* (23%), *HSP90AA1* (21%), *NFKB1A* (18%), *CALR* (17%), *IRS2* (13%), *PDPR* (12%), *MAP2K3* (12%) and *CCND3* (12%). Novel SNVs and CNVs were both more frequent in PD samples (Fig. 4A). KEGG analysis revealed more tumor related signaling pathways could be enriched by novel gene-level CNVs comparing with novel SNV. PI3K-Akt signaling pathway, chemical carcinogenesis-reactive oxygen species, Epstein-Barr virus infection and FoxO signaling pathway were the top enriched tumor related pathways (Fig. 4B). Pathways including platinum drug resistance and PD-L1 expression and PD-1 checkpoint pathway in cancer were also enriched.

### Impact of subsequent chemotherapy on genomic alterations of ctDNA driven by first-line treatment

Peripheral blood samples collected from 10 patients during treatments of subsequent regimens were analyzed. Both CNI and CNV burden values increased at second-line post-treatment relative to baseline and first-line PD. CCF and TMB values were not significantly different at these three time points (Fig. 5A). The number of genes with novel CNVs at second-line post-treatment was more than that at first-line PD (186 vs. 45). KEGG analysis of novel SNVs at second-line post-treatment showed that cAMP signaling pathway and Fanconi anemia pathway were the top enriched tumor related pathway (Fig. 5B).

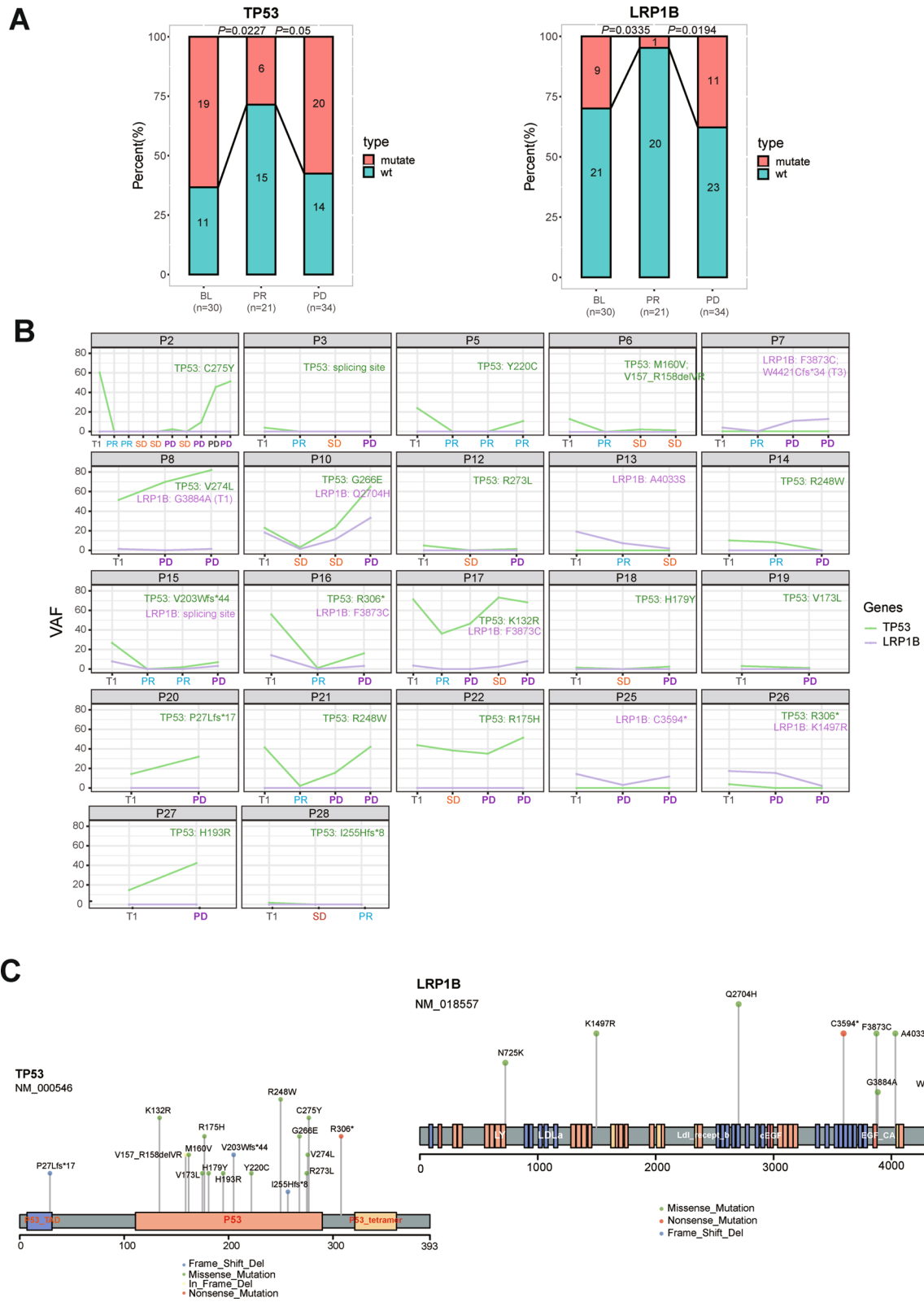
Multiple common signaling pathways enriched by novel gene-level CNVs between first-line PD and second-line post-treatment could be identified. NF-kappa B signaling pathway, platinum drug resistance pathway and FoxO signaling pathway could be enriched by novel CNVs at second-line post-treatment rather than at first-line PD (Fig. 5C).

### Discussion

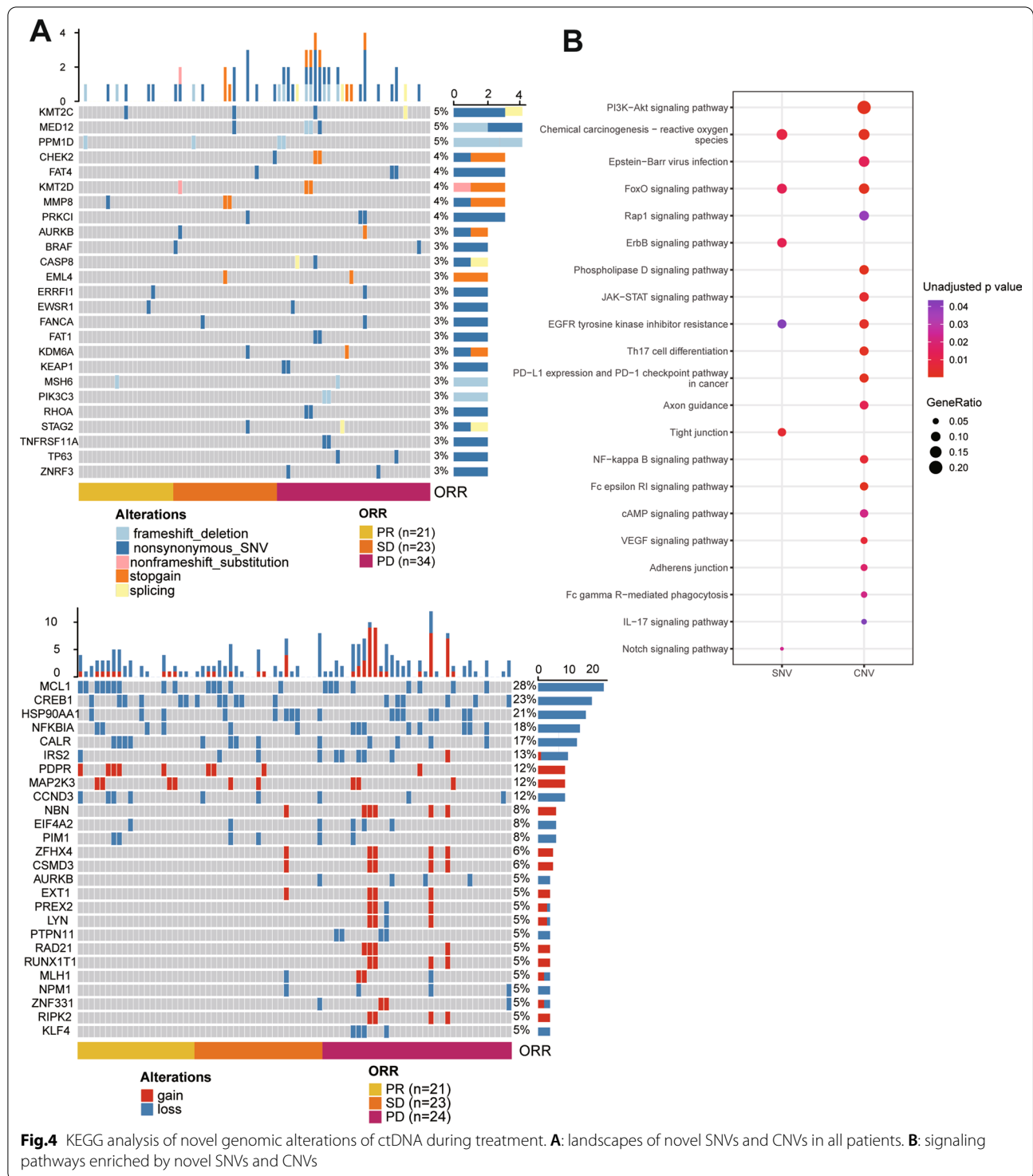
In the present study, both baseline and dynamic changes of genomic features in ctDNA were found to be correlated with the treatment efficacy of first-line chemotherapy for AGC patients. Fluctuating changes in SNVs and CNVs could be observed during sequential chemotherapy, where the impact of the subsequent regimens on the molecular changes driven by first-line therapy was determined to be subtle. Under the pressure of conventional chemotherapy, novel SNVs and CNVs could be enriched into multiple common oncologic signaling pathways, where their distributions were established to be highly heterogeneous among patients.

The correlation between the genomic features of ctDNA and the chemotherapy efficacy in AGC patients has been analyzed in previous studies, where most of the studies were focused on reporting the data of somatic mutations, including number of gene mutations and VAF [12, 14, 20]. Recent more detailed analyses of ctDNA revealed that the molecular changes correlated with anti-HER2 targeted therapy and anti-PD-1 immunotherapy [15, 16, 18]. Nevertheless, ctDNA analysis during conventional sequential chemotherapy—which is still the standard treatment pattern for most AGC patients in China—have not been fully explored.

We found that AGC patients who had higher baseline values of CNI and CNV burden in ctDNA were more likely to respond to platinum-based chemotherapy. Somatic CNVs were correlated with chromosome instability, which resulted from defects in mitosis and premitotic replication stress [21]. In AGC patients, it has been reported that high level chromosome instability was associated with sensitivity to platinum-based chemotherapy [22]. In colorectal cancer patients, tumors with increased copy number alterations also indicated sensitivity to chemotherapy plus bevacizumab [23]. Also, chromosome instability has been found to be associated with intrinsic resistance to taxanes [24]. In our study, we found that both CNI and CNV burden in ctDNA recovered to their baseline levels at treatment failure of first-line regimens. The responses to paclitaxel-based and irinotecan-based regimens as second-line treatments were both found to be poor in our cohort. This was in line with the findings of previous randomized controlled trials [9]. Our findings suggest that patients with high CNI



**Fig.3** Fluctuating changes of genomic alterations during sequential chemotherapy. **A:** changes of percentages of *TP53* and *LRP1B* mutations during treatment. **B:** fluctuating changes of *TP53* and *LRP1B* frequency during treatment. **C:** involved mutation sites of *TP53* and *LRP1B*



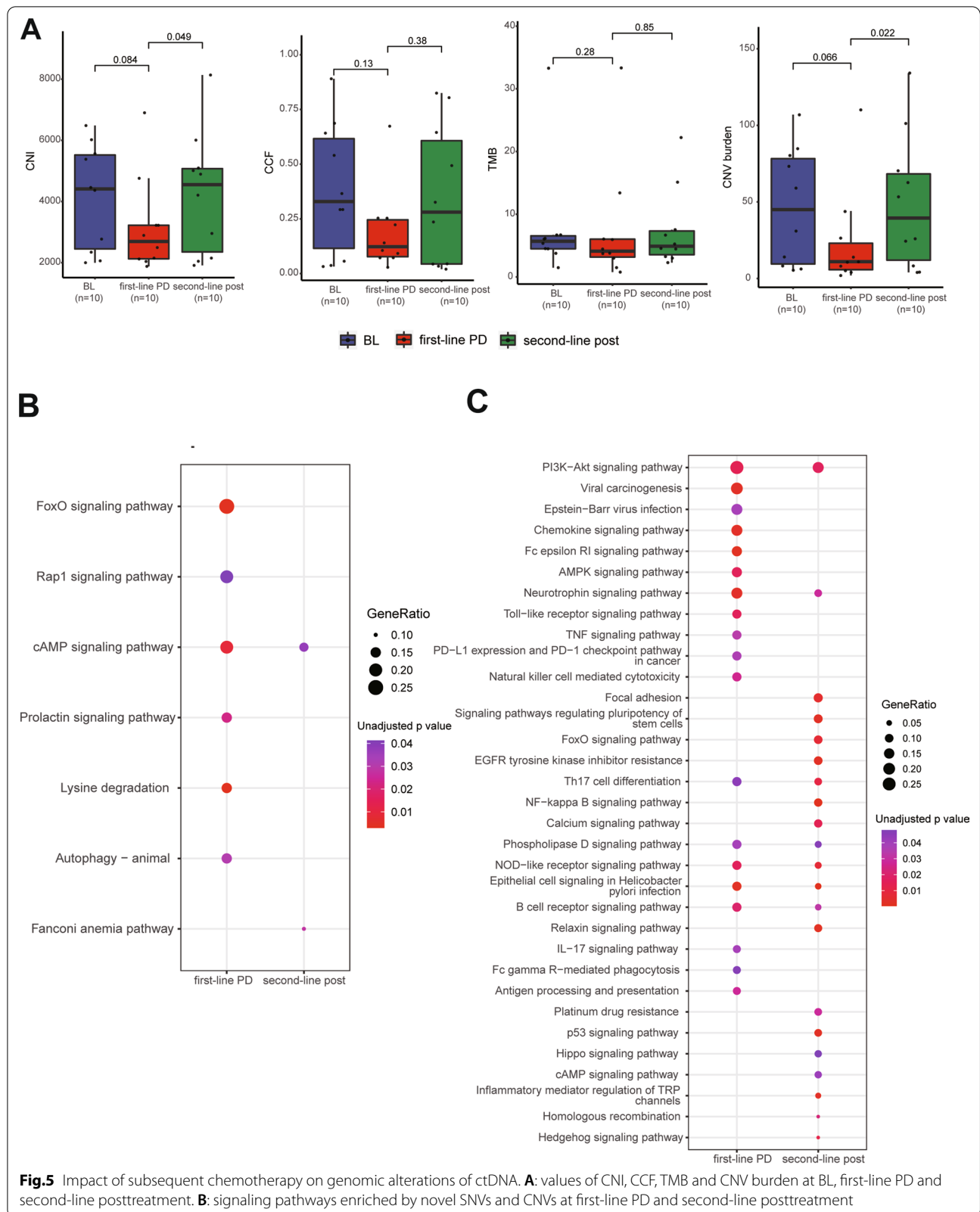
**Fig.4** KEGG analysis of novel genomic alterations of ctDNA during treatment. **A:** landscapes of novel SNVs and CNVs in all patients. **B:** signaling pathways enriched by novel SNVs and CNVs

and high CNV burden may not benefit from the standard pattern of subsequent chemotherapy of AGC.

Even though the values of CNI and CNV burden at disease progression were similar to those at baseline, treatment resistance still occurred. Novel SNVs and

CNVs were observed in ctDNA at PD. Compared with somatic mutations, platinum-based chemotherapy could cause more gene-level CNVs in ctDNA. A pan-cancer study demonstrated the existence of a positive linear influence of CNVs on the expression of most of





the genes, which was indicative of the presence of a direct effect of CNVs on gene transcriptional level [25]. Furthermore, gene-level CNVs are considered as common and important molecular alterations in tumors and are usually the upstream changes that trigger the cascade of signaling chaos [26]. In our study, KEGG analysis of novel CNVs showed that several oncologic and treatment-resistance pathways, including EGFR tyrosine kinase inhibitor resistance and platinum drug resistance, were enriched. These novel genomic alterations could be attributed to resistance to the first-line chemotherapy.

Tumor evolutionary theory provides an alternative strategy to optimize the current treatment paradigm of cancer [27]. Spontaneous fitness trade-offs are the cost of evolution during tumor adaptation to the drastic external pressure of chemotherapy. These fitness trade-offs lead to the appearance of vulnerability in tumor cells to specific subsequent therapies [28–30]. This phenomenon creates an opportunity for oncologists to anticipate tumor evolutionary process under the pressure of certain drugs and design an optimal subsequent regimen to prolong disease control [31]. However, although we found that genes with novel SNVs and CNVs could be identified during and after the whole course of the treatment, the corresponding distributions among AGC patients were highly heterogeneous.

Molecular matched therapy has been performed based on molecular alterations detected in tumors. However, the benefit of this strategy in subsequent treatment of gastric cancer is uncertain. SHIVA trial did not show a significant benefit, and the absolute time of survival extension in VIKTORY trial was reported to be only about 2 months [32, 33]. In our study, the genes involved in relevant signaling pathways were found to be highly heterogeneous among patients. It has been concluded that while common aberrant pathways might result in treatment failure, inter-patient heterogeneity and signaling crosstalk could lead to inefficacy of single-targeting molecular agents [34]. Thus, a multi-targeting strategy should be employed in subsequent regimens of AGC patients. In EPOC1706 trial, lenvatinib plus pembrolizumab in the first-line or the second-line setting of AGC patients showed an objective response rate of 69%, which was much higher than that of previously-employed standard second-line chemotherapy of AGC patients [35]. REGONIVO trial also showed a high response rate for the treatment of refractory AGC patients with regorafenib plus nivolumab [36]. The synergistic effects in place between lenvatinib/regorafenib and immune checkpoint inhibitors such as pembrolizumab and nivolumab, together with the fact that both lenvatinib and regorafenib are multi-targeting tyrosine

kinase inhibitors of VEGFR, FGFR, PDGFR, KIT and RET, should both account for such enhanced clinical efficacy [37, 38].

Tumor cells can adapt to external treatment pressure by acquisition of new genomic or epigenomic changes [39]. This means that the dynamics of tumor evolution should also be considered in devising therapeutic patterns for cancer. In a study on the genomic landscape of relapsed acute lymphoblastic leukemia, the prevalence of relapse-specific somatic alterations in patients showing very early relapse (<9 months from diagnosis) was found to be significantly less than that in the patients who exhibited an early or late relapse [40]. The median PFS of AGC patients in the first-line treatment is normally 6–8 months. Despite being for different kinds of tumors, these results suggest that gastric cancer cells might not have long enough time to acquire new driver somatic alterations under treatment pressure. In this context, Li et al. have also reported quite similar CNV profiles in paired tissue samples collected before and after neoadjuvant chemotherapy from gastric cancer patients who had no response [41].

The present study bears certain limitations. First, the fact that only 30 patients were enrolled in our study made the sample size relatively small. Nevertheless, analysis of 108 ctDNA samples could present a representative view over the dynamic molecular changes during the treatment. Second, analysis of subsequent therapy was only performed in 10 patients. The fact that the available information found in the literature regarding the dynamic molecular changes during sequential chemotherapy of AGC patients is limited draws attention to the importance of the results of our study. Third, the analysis of SNVs and gene-level CNVs in our study was not accompanied by the examination of the possible correlations between them and the gene expression. Thus, it is our aim to investigate the transcriptional data in future studies using a larger sample size.

In summary, baseline and dynamic changes of genomic features of ctDNA could be biomarkers to predict treatment efficacy of platinum-based first-line chemotherapy in AGC patients. After treatment with standard chemotherapy regimens, convergent oncologic pathway enrichment yet characterized by inter-patient heterogeneous gene-level CNVs was identified, which provided valuable information for the optimization of the sequential treatment of AGC patients.

## Materials and methods

### Patients and samples

Peripheral blood samples were collected from gastric cancer patients who were treated in Department of Oncology, Ruijin Hospital and met the following criteria:

(1) age over 18 years; (2) histologically confirmed HER2 negative stomach adenocarcinoma; (3) treatment naïve locally advanced, recurrent or metastasis disease; (4) adequate performance status and organ function to receive standard chemotherapy. Written informed consent was provided by all patients before collecting the peripheral blood samples. The protocol was approved by ethics committee of Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, People's Republic of China.

Blood samples were collected at times of baseline, efficacy assessment and disease progression. Samples collected at treatment failure of one regimen were considered as baseline samples for the regimen that followed. Regimens with at least two samples during treatment were eligible for further analysis. Treatment efficacy was assessed by radiographic examination (contrast enhanced CT or MRI) following the routine clinical protocols every 8–9 weeks. Objective response was assessed according to the RECIST criteria (v1.1). Clinicopathological information of all the patients was recorded, and PFS and OS were followed up. The best objective response during first-line treatment was used to stratify patients into R and NR groups. During the whole course of treatment, treatment response at each sampling timepoint was compared with last timepoint according to RECIST criteria (v1.1). Tumor response of each timepoint was listed in Additional file 11: Table S5.

#### DNA extraction, sequencing and data processing

Plasma and blood cells were separated by centrifugation at 1600 g for 10 min. Cell-free DNA (cfDNA) was isolated from plasma using MagMAX™ Cell-Free DNA Isolation Kit (Thermo Fisher Scientific, Waltham, MA, USA). Blood cell DNA was extracted from blood cells with TIANamp Blood DNA Kit (TIANGEN, Beijing, China). The concentration of DNA was measured by Qubit® dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA), and the quality of DNA was assessed by Agilent 2100 BioAnalyzer (Agilent, USA). Genomic DNA was sheared into 150–200 base-pair (bp) fragments with Covaris M220 Focused-ultrasonicator Instrument (Covaris, Massachusetts, USA). Fragmented DNA was used to construct the library using KAPA Hyper Preparation Kit (Kapa Biosystems, USA) according to the manufacturer's instruction. DNA was hybridized to one in-house panel (Genecast, Wuxi, China) covering 2.43 Mb of the human genome with 1632 genes (Additional file 12: Table S6). The captured library was sequenced on Illumina NovaSeq 6000 according to the manufacturer's instruction, producing paired-end reads with the length of each end as 150 bp.

Preliminary sequencing results in BCL format were converted to FASTQ files using bcl2fastq (v2.20.0). FASTQ format reads were processed using Trimmomatic (v0.39) [42] for adapter trimming and low-quality reads filtering. Processed reads were mapped to the reference genome (hg19) using BWA (0.7.17) [43]. Mapped results were sorted and marked for duplications using Picard toolkit (version 2.1.0, <https://broadinstitute.github.io/picard/>), and then were realigned using GATK (version 3.7) [44]. The quality control data were shown in Additional file 13: Table S7. The average sequencing depth of CF and BC samples was  $>1000\times$  and  $>200\times$ , respectively. The distribution of fragment sizes for each sample is shown in Additional file 4: Figure S4.

#### Somatic SNV calling

SNVs were called via VarDict (v1.5.1) and FreeBayes (v1.2.0) from processed mapping results in pair of tumor and control samples, respectively. The candidate somatic mutations were annotated with ANNOVAR (2015Jun17) [45] and then were filtered by ExAC (<http://exac.broadinstitute.org>), gnomAD (<http://gnomad-sg.org/>), COSMIC (<https://cancer.sanger.ac.uk/cosmic>), dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>) databases. Nonsynonymous mutations among the exonic and splicing regions were kept for the final somatic mutation data set (Additional file 5: Figure S5). The filter parameters of SNV include the following criteria: (1) exclude mutations located in intergenic regions or intronic regions. (2) exclude blacklisted mutations. (3) exclude mutations with sequencing depth  $<120$ . (4) exclude mutations with allele frequency  $<1\%$ . (5) exclude mutations with reads  $<5$ . (6) exclude synonymous mutations. (7) exclude allele frequency  $\geq 0.002$  in the Exome Aggregation Consortium (ExAC) database. (8) exclude mutations also present in blood control samples and the frequency in tumor sample is less than five times frequency in blood control sample (Additional file 14: Table S8).

#### Somatic CNV calling

Taking blood cell sample as the control, somatic CNVs were called from plasma sample for each gene included in the captured panel using CNVkit (0.9.5.dev0) with batch mode [46]. In the process of calling, biases of read depth, GC content and repetitive sequence were removed through normalization (Additional file 6: Figure S6). The filtering criterion of CNV is the number of capture region more than 5, and the threshold of gain and loss are copy ratio  $>2.5$  and copy ratio  $<1.5$ , respectively. The CNV value was defined as  $\log_2$  transformed ratio of normalized read depth on each gene between plasma and blood cell samples (Additional file 15: Table S9).

### Calculation of tumor mutation burden (TMB)

TMB were determined by the nonsynonymous somatic mutations among the exonic and splicing regions. Alterations likely or known to be bona fide oncogenic drivers were excluded (Additional file 7: Figure S7). TMB per megabase were calculated with the total number of mutations divided by the total bases of the target panel with  $\geq 500$  bp coverage.

### Definition of meanVAF and maxVAF

The maxVAF value equals to the maximum VAF of all somatic mutations in each sample. The meanVAF value was defined as the mean value of all mutations' VAFs in each sample.

### Calculation of CNV burden

CNV burden is equal to the total number of CNVs in each sample. Considering the two types of CNVs, the CNV burden was divided into gain and loss. The total CNV of gain and total CNV of loss were defined as the CNV gain burden and CNV loss burden, respectively.

### Estimation of ctDNA content fraction (CCF)

CCF of plasma samples was estimated by a maximum likelihood model based on SNVs and CNVs in the paired blood cell and plasma samples. Somatic and germline SNPs that met the following criteria were used to build the model: (1) with a minimum depth of  $50 \times$  in the paired samples; (2) not on genes with high polymorphism; (3) no InDels in the 50 bp upstream or downstream regions; (4) not in a copy number gain region; (5) germline SNPs with significantly different variant allele frequencies (VAFs) in the paired samples, or somatic SNPs with VAFs significantly higher than background noise. These SNPs were defined as informative SNPs and were clustered into multiple groups according to their VAFs, local copy numbers and hypothetical genotypes. The hypothetical genotypes in cfDNA and ctDNA were determined by the VAFs in the paired samples and the copy number in the plasma sample. Each cluster represents a unique ctDNA source. We then calculated the likelihood of observing SNPs under given CCFs in each cluster. By maximizing the likelihood, CCF of each cluster could therefore be estimated. Cluster with the highest CCF was from the main source of ctDNA, and its CCF was the output of the final estimation.

### Estimation of copy number instability (CNI)

After correction for GC content and length of target region using proprietary algorithms for each region, the read counts were transformed into log<sub>2</sub> ratios and were

converted into Z-score based on Gaussian transformations versus a normal control group ( $n = 30$ ). The target regions that satisfied the Z-score greater than the 95th percentile plus twice-times absolute standard deviation of the normal control group were retained, and these Z-scores were summed as the CNI score.

### Statistical analysis

For the comparison of continuous variables, we chose Wilcoxon rank sum test. When a continuous variable was transformed into a categorical one, the best cutoff was determined through pROC (1.18.0) package [47]. Comparisons between categorical variables were conducted using Fisher's exact test. Kaplan–Meier survival curves with log-rank test were generated by survminer (0.4.9) and survival packages (3.3–1). The visualization of gene alteration landscapes was profiled by ComplexHeatmap package (2.12.0) [48]. The KEGG analysis of newly emerged gene alterations was performed on clusterProfiler package (v4.4.4) [49]. We customized 1632 panel genes as background gene sets for pathways enrichment. All analyses were carried on R (4.2.1) program.  $P < 0.05$  was considered as statistically significant.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-022-03567-5>.

**Additional file 1: Figure S1.** Baseline genomic features of ctDNA and treatment response of first-line therapy. A: values of meanVAF, maxVAF, CNV gain burden and CNV loss burden in R and NR groups. B: cut-off value of CNI and its correlation with treatment response. C: correlation between CNI value and patients' overall survival.

**Additional file 2: Figure S2.** Dynamic change of genomic features of ctDNA and landscapes of SNVs and CNVs. A: values of meanVAF, maxVAF, CNV gain burden, CNV loss burden at baseline, disease progression and during treatment. B: landscapes of SNVs of all samples. C: landscapes of CNVs of all samples.

**Additional file 3: Figure S3.** Correlation between dynamic changes of genomic features of ctDNA and progression free survival of first-line therapy.

**Additional file 4: Figure S4.** Insert size distribution.

**Additional file 5: Figure S5.** Coverage distribution.

**Additional file 6: Figure S6.** All uniform Manhattan of blood control samples.

**Additional file 7: Figure S7.** All uniform Manhattan of cell-free samples

**Additional file 8: Table S1.** Treatment paradigms of 30 gastric cancer and the corresponding clinical outcomes. **Table S2.** Examination of the differences between the clinical characteristics of patients in NR and R groups at first-line treatment.

**Additional file 9: Table S3.** SNV information.

**Additional file 10: Table S4.** CNV values of all samples.

**Additional file 11: Table S5.** Information of sampling timepoints and treatment response.

**Additional file 12: Table S6.** Gene list of NGS panel.

**Additional file 13: Table S7.** Quality control.

**Additional file 14: Table S8.** SNV landscape of all ctDNA samples.

**Additional file 15: Table S9.** CNV landscape of all ctDNA samples.

### Acknowledgements

The study was supported by National Science Foundation of China (81972707 and 81802319) and the Collaborative Innovation Cluster Project of Shanghai Municipal Health Commission (2020CXJQ03).

### Author contributions

ZJ and XWQ conceived and designed the research. XWQ, ZCF, JJ, MT, WJW and SGCF contributed to the collection of specimens, radiological evaluation, and discussion of clinical significance. ZJ and ZCF contributed to design of the analysis. XF, SDB, and CYW led the data analysis. XF, SDB and ZCF contributed to the data analysis. ZJ, XWQ and ZCF wrote the manuscript. ZJ supervised the whole project. All authors read and approved the final manuscript.

### Declarations

#### Competing interests

The authors declared that no competing interest exists.

#### Author details

<sup>1</sup>Department of Oncology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, No. 197 Ruijin er Road, Shanghai 200025, China. <sup>2</sup>Department of Oncology, Wuxi Branch of Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, No 197 Zhixian Road, Xinwu District, Wuxi 214028, China. <sup>3</sup>Shanghai Institute of Digestive Surgery, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, No. 197 Ruijin er Road, Shanghai 200025, China. <sup>4</sup>Genecast Biotechnology Co., Ltd, Wuxi City 214104, Jiangsu, China. <sup>5</sup>Department of Radiation Oncology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, No. 197 Ruijin er Road, Shanghai 200025, China. <sup>6</sup>State Key Laboratory of Oncogenes and Related Genes, Shanghai Jiao Tong University, Shanghai 200025, China.

Received: 9 May 2022 Accepted: 2 August 2022

Published online: 12 August 2022

### References

- Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ, He J. Cancer statistics in China. *CA Cancer J Clin*. 2016;66(2):115–32.
- Gao K, Wu J. National trend of gastric cancer mortality in China (2003–2015): a population-based study. *Cancer Commun (Lond)*. 2019;39(1):24.
- Joshi SS, Badgwell BD. Current treatment and recent progress in gastric cancer. *CA Cancer J Clin*. 2021;71(3):264–79.
- Nagaraja AK, Kikuchi O, Bass AJ. Genomics and targeted therapies in gastroesophageal adenocarcinoma. *Cancer Discov*. 2019;9(12):1656–72.
- Oh SC, Sohn BH, Cheong JH, Kim SB, Lee JE, Park KC, Lee SH, Park JI, Park YY, Lee HS, et al. Clinical and genomic landscape of gastric cancer with a mesenchymal phenotype. *Nat Commun*. 2018;9(1):1777.
- Sa JK, Hong JY, Lee IK, Kim JS, Sim MH, Kim HJ, An JY, Sohn TS, Lee JH, Bae JM, et al. Comprehensive pharmacogenomic characterization of gastric cancer. *Genome Med*. 2020;12(1):17.
- Cunningham D, Starling N, Rao S, Iveson T, Nicolson M, Coxon F, Middleton G, Daniel F, Oates J, Norman AR, et al. Capecitabine and oxaliplatin for advanced esophagogastric cancer. *N Engl J Med*. 2008;358(1):36–46.
- Ajani JA, Rodriguez W, Bodoky G, Moiseyenko V, Lichinitser M, Gorbunova V, Vynnychenko I, Garin A, Lang I, Falcon S. Multicenter phase III comparison of cisplatin/S-1 with cisplatin/infusional fluorouracil in advanced gastric or gastroesophageal adenocarcinoma study: the FLAGS trial. *J Clin Oncol*. 2010;28(9):1547–53.
- Hironaka S, Ueda S, Yasui H, Nishina T, Tsuda M, Tsumura T, Sugimoto N, Shimodaira H, Tokunaga S, Moriwaki T, et al. Randomized, open-label, phase III study comparing irinotecan with paclitaxel in patients with advanced gastric cancer without severe peritoneal metastasis after failure of prior combination chemotherapy using fluoropyrimidine plus platinum: WJOG 4007 trial. *J Clin Oncol*. 2013;31(35):4438–44.
- Kilgour E, Rothwell DG, Brady G, Dive C. Liquid biopsy-based biomarkers of treatment response and resistance. *Cancer Cell*. 2020;37(4):485–95.
- Alix-Panabieres C, Pantel K. Liquid biopsy: from discovery to clinical application. *Cancer Discov*. 2021;11(4):858–73.
- Wu R, Shi C, Chen Q, Wu F, Li Q. Detection of circulating tumor cell DNA for monitoring advanced gastric cancer. *Int J Clin Exp Pathol*. 2020;13(2):203–11.
- Lengyel CG, Hussain S, Trapani D, El Bairi K, Altuna SC, Seeber A, Odhiambo A, Habeeb BS, Seid F. The emerging role of liquid biopsy in gastric cancer. *J Clin Med*. 2021;10(10):2108.
- Parikh AR, Mojtahed A, Schneider JL, Kanter K, Van Seventer EE, Fetter IJ, Thabet A, Fish MG, Teshome B, Fosbenner K, et al. Serial ctDNA monitoring to predict response to systemic therapy in metastatic gastrointestinal cancers. *Clin Cancer Res*. 2020;26(8):1877–85.
- Wang Y, Zhao C, Chang L, Jia R, Liu R, Zhang Y, Gao X, Li J, Chen R, Xia X, et al. Circulating tumor DNA analyses predict progressive disease and indicate trastuzumab-resistant mechanism in advanced gastric cancer. *EBioMedicine*. 2019;43:261–9.
- Jin Y, Chen DL, Wang F, Yang CP, Chen XX, You JQ, Huang JS, Shao Y, Zhu DQ, Ouyang YM, et al. The predicting role of circulating tumor DNA landscape in gastric cancer patients treated with immune checkpoint inhibitors. *Mol Cancer*. 2020;19(1):154.
- Zhang M, Qi C, Wang Z, Chen H, Zhao X, Zhang X, Zhou Y, Gao C, Bai Y, Jia S, et al. Molecular characterization of ctDNA from Chinese patients with advanced gastric adenocarcinoma reveals actionable alterations for targeted and immune therapy. *J Mol Med (Berl)*. 2021;99(9):1311–21.
- Zhang C, Chen Z, Chong X, Chen Y, Wang Z, Yu R, Sun T, Chen X, Shao Y, Zhang X, et al. Clinical implications of plasma ctDNA features and dynamics in gastric cancer treated with HER2-targeted therapies. *Clin Transl Med*. 2020;10(8):e254.
- Zhou C, Ma T, Shi M, Xi W, Wu J, Yang C, Zhu Z, Zhang J. Dose-finding study of modified FLOT (mFLOT) regimen as first-line treatment in Chinese patients with metastatic adenocarcinoma of stomach. *Cancer Chemother Pharmacol*. 2020;85(1):113–9.
- Normando SRC, Delgado PO, Rodrigues A, David Filho WJ, Fonseca FLA, Cruz F, Del Giglio A. Circulating free plasma tumor DNA in patients with advanced gastric cancer receiving systemic chemotherapy. *BMC Clin Pathol*. 2018;18:12.
- Burrell RA, McClelland SE, Endesfelder D, Groth P, Weller MC, Shaikh N, Domingo E, Kanu N, Dewhurst SM, Gronroos E, et al. Replication stress links structural and numerical cancer chromosomal instability. *Nature*. 2013;494(7438):492–6.
- Ott K, Vogelsang H, Mueller J, Becker K, Muller M, Fink U, Siewert JR, Hofler H, Keller G. Chromosomal instability rather than p53 mutation is associated with response to neoadjuvant cisplatin-based chemotherapy in gastric carcinoma. *Clin Cancer Res*. 2003;9(6):2307–15.
- Smeets D, Miller IS, O'Connor DP, Das S, Moran B, Boeckx B, Gaiser T, Betge J, Barat A, Klingler R, et al. Copy number load predicts outcome of metastatic colorectal cancer patients receiving bevacizumab combination therapy. *Nat Commun*. 2018;9(1):4112.
- McClelland SE, Burrell RA, Swanton C. Chromosomal instability: a composite phenotype that influences sensitivity to chemotherapy. *Cell Cycle*. 2009;8(20):3262–6.
- Shao X, Lv N, Liao J, Long J, Xue R, Ai N, Xu D, Fan X. Copy number variation is highly correlated with differential gene expression: a pan-cancer study. *BMC Med Genet*. 2019;20(1):175.
- Hehir-Kwa JY, Pfundt R, Veltman JA. Exome sequencing and whole genome sequencing for the detection of copy number variation. *Expert Rev Mol Diagn*. 2015;15(8):1023–32.
- Enriquez-Navas PM, Wojtkowiak JW, Gatenby RA. Application of evolutionary principles to cancer therapy. *Can Res*. 2015;75(22):4675–80.
- Lin KH, Rutter JC, Xie A, Pardieu B, Winn ET, Bello RD, Forget A, Itzykson R, Ahn YR, Dai Z, et al. Using antagonistic pleiotropy to design a chemotherapy-induced evolutionary trap to target drug resistance in cancer. *Nat Genet*. 2020;52(4):408–17.
- Hausser J, Alon U. Tumour heterogeneity and the evolutionary trade-offs of cancer. *Nat Rev Cancer*. 2020;20(4):247–57.
- Wang L, Leite de Oliveira R, Huijberts S, Bosdriesz E, Pencheva N, Brunen D, Bosma A, Song JY, Zevenhoven J, Los-de Vries GT, et al. An acquired vulnerability of drug-resistant melanoma with therapeutic potential. *Cell*. 2018;173(6):1413–25.



31. Stankova K, Brown JS, Dalton WS, Gatenby RA. Optimizing cancer treatment using game theory: a review. *JAMA Oncol.* 2019;5(1):96–103.
32. Le Tourneau C, Delord JP, Goncalves A, Gavoille C, Dubot C, Isambert N, Campone M, Tredan O, Massiani MA, Mauborgne C, et al. Molecularly targeted therapy based on tumour molecular profiling versus conventional therapy for advanced cancer (SHIVA): a multicentre, open-label, proof-of-concept, randomised, controlled phase 2 trial. *Lancet Oncol.* 2015;16(13):1324–34.
33. Lee J, Kim ST, Kim K, Lee H, Kozarewa I, Mortimer PGS, Odegaard JI, Harrington EA, Lee J, Lee T, et al. Tumor genomic profiling guides patients with metastatic gastric cancer to targeted treatment: the VIKTORY umbrella trial. *Cancer Discov.* 2019;9(10):1388–405.
34. Boumahdi S, de Sauvage FJ. The great escape: tumour cell plasticity in resistance to targeted therapy. *Nat Rev Drug Discovery.* 2020;19(1):39–56.
35. Kawazoe A, Fukuoka S, Nakamura Y, Kuboki Y, Wakabayashi M, Nomura S, Mikamoto Y, Shima H, Fujishiro N, Higuchi T, et al. Lenvatinib plus pembrolizumab in patients with advanced gastric cancer in the first-line or second-line setting (EPOC1706): an open-label, single-arm, phase 2 trial. *Lancet Oncol.* 2020;21(8):1057–65.
36. Fukuoka S, Hara H, Takahashi N, Kojima T, Kawazoe A, Asayama M, Yoshii T, Kotani D, Tamura H, Mikamoto Y, et al. Regorafenib plus nivolumab in patients with advanced gastric or colorectal cancer: an open-label, dose-escalation, and dose-expansion phase Ib trial (REGONIVO, EPOC1603). *J Clin Oncol.* 2020;38(18):2053–61.
37. Li J, Qin S, Xu R, Yau TC, Ma B, Pan H, Xu J, Bai Y, Chi Y, Wang L, et al. Regorafenib plus best supportive care versus placebo plus best supportive care in Asian patients with previously treated metastatic colorectal cancer (CONCUR): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol.* 2015;16(6):619–29.
38. Al-Salama ZT, Syed YY, Scott LJ. Lenvatinib: a review in hepatocellular carcinoma. *Drugs.* 2019;79(6):665–74.
39. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell.* 2017;168(4):613–28.
40. Li B, Brady SW, Ma X, Shen S, Zhang Y, Li Y, Szlachta K, Dong L, Liu Y, Yang F, et al. Therapy-induced mutations drive the genomic landscape of relapsed acute lymphoblastic leukemia. *Blood.* 2020;135(1):41–55.
41. Li Z, Gao X, Peng X, May Chen MJ, Li Z, Wei B, Wen X, Wei B, Dong Y, Bu Z, et al. Multi-omics characterization of molecular features of gastric cancer correlated with response to neoadjuvant chemotherapy. *Sci Adv.* 2020;6(9):eaay4211.
42. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–20.
43. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics.* 2009;25(14):1754–60.
44. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, et al. From FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics.* 2013;43:11 10 11–11 10 33.
45. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16):e164.
46. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput Biol.* 2016;12(4):e1004873.
47. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* 2011;12:77.
48. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics.* 2016;32(18):2847–9.
49. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Camb).* 2021;2(3):100141.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

