ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ESMO OPEN SCIENCE FOR OPTIMAL CANCER CARE

ORIGINAL RESEARCH

Sequencing paired tumor DNA and white blood cells improves circulating tumor DNA tracking and detects pathogenic germline variants in localized colon cancer☆

F. Gimeno-Valiente[1†], J. Martín-Arana[2,3†], R. Tébar-Martínez[2], V. Gambardella[2], C. Martínez-Ciarpaglini[3,4], B. García-Micó[2,3], B. Martínez-Castedo[2,3], B. Palomar[4], M. García-Bartolomé[2], V. Seguí[2], M. Huerta[2], D. Moro-Valdezate[5], V. Pla-Martí[5], L. Pérez-Santiago[5], S. Roselló[2,3], D. Roda[2,3], A. Cervantes[2,3*‡] & N. Tarazona[2,3*‡]

[1]Cancer Evolution and Genome Instability Laboratory, University College London Cancer Institute, London, UK; [2]Department of Medical Oncology, INCLIVA Biomedical Research Institute, University of Valencia, Valencia; [3]CIBERONC, Instituto de Salud Carlos III, Madrid; [4]Department of Pathology, INCLIVA Biomedical Research Institute, University of Valencia, Valencia; [5]Colorectal Surgery Unit, INCLIVA Biomedical Research Institute, Hospital Clínico Universitario de Valencia, Department of Surgery, University of Valencia, Valencia, Spain

Available online 10 November 2023

Background: In the setting of localized colon cancer (CC), circulating tumor DNA (ctDNA) monitoring in plasma has shown potential for detecting minimal residual disease (MRD) and predicting a higher risk of recurrence. With the tumor-only sequencing approach, however, germline variants may be misidentified as somatic variations, precluding the possibility of tracking in up to 11% of patients due to a lack of known somatic mutations. In this study, we assess the potential value of adding white blood cells (WBCs) to tumor tissue sequencing to enhance the accuracy of sequencing results.

Patients and Methods: A total of 148 patients diagnosed with localized CC were prospectively recruited at the Hospital Clínico Universitario in Valencia (Spain). Employing a custom 29-gene panel, sequencing was conducted on tumor tissue, plasma and corresponding WBCs. Droplet digital PCR and amplicon-based NGS were performed on plasma samples post-surgery to track MRD. Oncogenic somatic variants were identified by annotating with COSMIC, OncoKB and an internal repository of pathogenic mutations database. A variant prioritization analysis, mainly characterized by the match of oncogenic mutations with the evidence levels defined in OncoKB, was carried out to select specific targeted therapies.

Results: Utilizing paired tumor and WBCs sequencing, we identified somatic mutations in all patients (100%) within our cohort, compared to 89% using only tumor tissue. Consequently, the top 10 most frequently mutated genes for plasma monitoring were altered. The sequencing of WBCs identified 9% of patients with pathogenic mutations in the germline, with *APC* and *TP53* being the most frequently mutated genes. Additionally, mutations in genes related to clonal hematopoiesis of indeterminate potential were detected in 27% of the cohort, with *TP53*, *KRAS*, and *KMT2C* being the most frequently altered genes. There were no observed differences in the sensitivity of monitoring MRD using ddPCR or amplicon-based NGS ($p = 1$). Ultimately, 41% of the patients harbored potentially targetable alterations at diagnosis.

Conclusion: The germline testing method not only enhanced sequencing results and raised the proportion of patients eligible for plasma monitoring, but also uncovered the existence of pathogenic germline variations, thereby aiding in the identification of patients at a higher risk of hereditary cancer syndromes.

Key words: colon cancer, circulating tumor DNA, precision medicine, germline variants, minimal residual disease

*Correspondence to: Prof. Andrés Cervantes, Avenida Blasco Ibañez, 17, 46010, Valencia, Spain. Tel: (+34)961973531
E-mail: andres.cervantes@uv.es (A. Cervantes).
*Dr Noelia Tarazona, Avenida Blasco Ibañez, 17, 46010, Valencia, Spain. Tel: +(34)961973531
E-mail: noetalla@incliva.es (N. Tarazona).

†These authors contributed equally as first authors.
‡These authors contributed equally as senior/corresponding author.
☆Note: Part of the content presented orally at the virtual ESMO Congress in 2020.

INTRODUCTION

Colon cancer (CC) is a heterogeneous disease representing the second leading cause of cancer-related death worldwide.[1] In localized CC, surgical pathology analysis of the surgical specimen provides the best estimation of the risk of relapse and guides adjuvant chemotherapy treatment (ACT) decision making. However, some patients might be mistreated with ACT if only TNM staging is considered, suggesting a need for more accurate prognostic and predictive

biomarkers.[2] In recent years, several studies have shown that circulating tumor DNA (ctDNA) detection immediately after surgery in patients diagnosed with localized CC provides direct evidence of minimal residual disease (MRD) and identifies those at high risk of recurrence.[3-12] Findings from the phase II randomized DYNAMIC study suggest that ctDNA-guided management could help improve clinician selection of patients with stage II CC who may benefit from ACT, as well as avoid overtreatment in those who do not present ctDNA.[13]

However, one of the most important issues in monitoring MRD is identifying genomic aberrations to be serially tracked in plasma. In our previous study,[6,10] up to 11% of patients with localized CC could not be monitored in plasma due to the lack of pathological somatic mutations in tumor samples. Several studies showed that sequencing artifacts, germline variants, and clonal hematopoiesis may confound the interpretation of sequencing results and complicate subsequent disease monitoring.[14] Accurately detecting tumor-derived mutations in ctDNA requires a combination of integrated digital error suppression approaches such as unique molecular identifiers, appropriate variant calling, multigene analysis, and high-depth sequencing.[15] In addition, the ctDNA fraction is extremely low in many tumors,[16] which may lead to false negatives in ∼18% of cases.[8] ctDNA detection methods should be highly sensitive and specific. Sequencing tumor DNA and matched white blood cells (WBCs) may improve sequencing analysis interpretation by filtering alterations and distinguishing ctDNA alterations from cell-free DNA (cfDNA) variants related to clonal hematopoiesis of indeterminate potential (CHIP). In this study, we aimed to track not only pathogenic variants but also any somatic mutations present, to increase the number of candidates for monitoring in plasma and therefore improve sensitivity in detecting MRD. Moreover, this new germline approach would enable us to identify not only CHIP-related mutations but also inherited genetic variations with clinical relevance for patients.

METHODS

Study population and sample collection

We recruited 148 patients diagnosed with localized CC at the Hospital Clínico Universitario in Valencia, Spain, between October 2015 and October 2019.[10] Tumor tissue, plasma, and WBCs were collected at diagnosis (pretreatment) for deep targeted sequencing (Supplementary Figure S1, available at https://doi.org/10.1016/j.esmoop.2023.102051). Tissue DNA was extracted from the primary tumor surgical specimen using the AllPrep DNA/RNA FFPE Kit (Qiagen) and sequenced with the QIAseq Targeted DNA Panel for Illumina Instruments protocol (May 2017; Qiagen) to identify specific mutations for each tumor. Germline DNA (gDNA) from 1 ml of the buffy coat was processed with chemagic DNA Blood 1 ml (Perkin-Elmer) for the chemagic instrument. cfDNA was extracted from 4 ml plasma samples using the QIAamp Circulating Nucleic Acid Kit (Qiagen). All procedures followed the manufacturer's instructions.

DNA concentration was measured using the Qubit fluorometer. DNA fragment lengths were assessed using the Agilent 4200 TapeStation System.

Samples and data from patients included in this study were provided by the INCLIVA Biobank (PT20/00029; B.0000768 ISCIII), part of the Valencian Biobanking Network and the Spanish National Biobanks Network and they were processed following standard operating procedures with the appropriate approval of the research and ethics committees. All participants provided written informed consent.

Library generation and sequencing

Libraries for gDNA, tumor DNA and cfDNA were prepared using the QIAseq Targeted DNA Panel kit from Qiagen, using a previously described customized CRC gene panel.[10] The manufacturer's protocol was followed, starting with an input of 40 ng for gDNA, 100 ng for tumor DNA and variable for cfDNA. Notably, the Universal PCR step comprised 18 cycles for gDNA, 20 cycles for tumor DNA and 17 cycles for cfDNA. The resulting libraries were assessed using the High Sensitivity D1000 DNA ScreenTape from Agilent on the TapeStation 4200 System. Equimolar volumes of libraries from the same DNA type (gDNA, tumor DNA or cfDNA) were multiplexed and diluted for loading and sequencing on the NextSeq instrument from Illumina, following the loading instructions provided by Qiagen. As a quality control measure for the entire variant determination process, samples NA12877 and NA12878 from the Coriell Institute were selected.

FASTQ preprocessing, quality control, and read mapping

Supplementary Figure S2, available at https://doi.org/10.1016/j.esmoop.2023.102051 illustrates the bioinformatics pipeline used in this study. The quality control of raw samples in FASTQ format was conducted using FastQC1 v0.11.8.[17] Adapter removal was carried out using Cutadapt2 v2.10,[18] and reads with a mean quality below Q30 were discarded using PRINSEQ v0.20.4.[19] Sequencing reads were aligned to the GRCh38.p12 human reference genome using BWA-mem v0.7.17.[20] The fgbio consensus pipeline (v1.1.0) was implemented for UMI extraction, grouping, and deduplication of reads from the same UMI family. Subsequently, BAM postprocessing was carried out using PICARD v2.18.6[21] and GATK v4.2.0.0.[22]

Variant calling and somatic variant prioritization

Variant calling of the primary tumor, plasma, and normal samples was conducted by combining the outputs from smCounter (v2),[23] LoFreq v2.1.5,[24] and Mutect2 (GATK v4.2.0.0) specifically for INDELs. For primary tumor variants, a minimum variant allele frequency (VAF) threshold of 5% was applied, while for plasma samples it was set at 0.01%. The resulting variants were then annotated using Variant Effect Predictor (Ensembl v102).[25] To avoid false positives, variants associated with CHIP were removed from the plasma samples and analyzed separately. We also carried out a manual review and curation of the candidate somatic mutations detected in each sample.

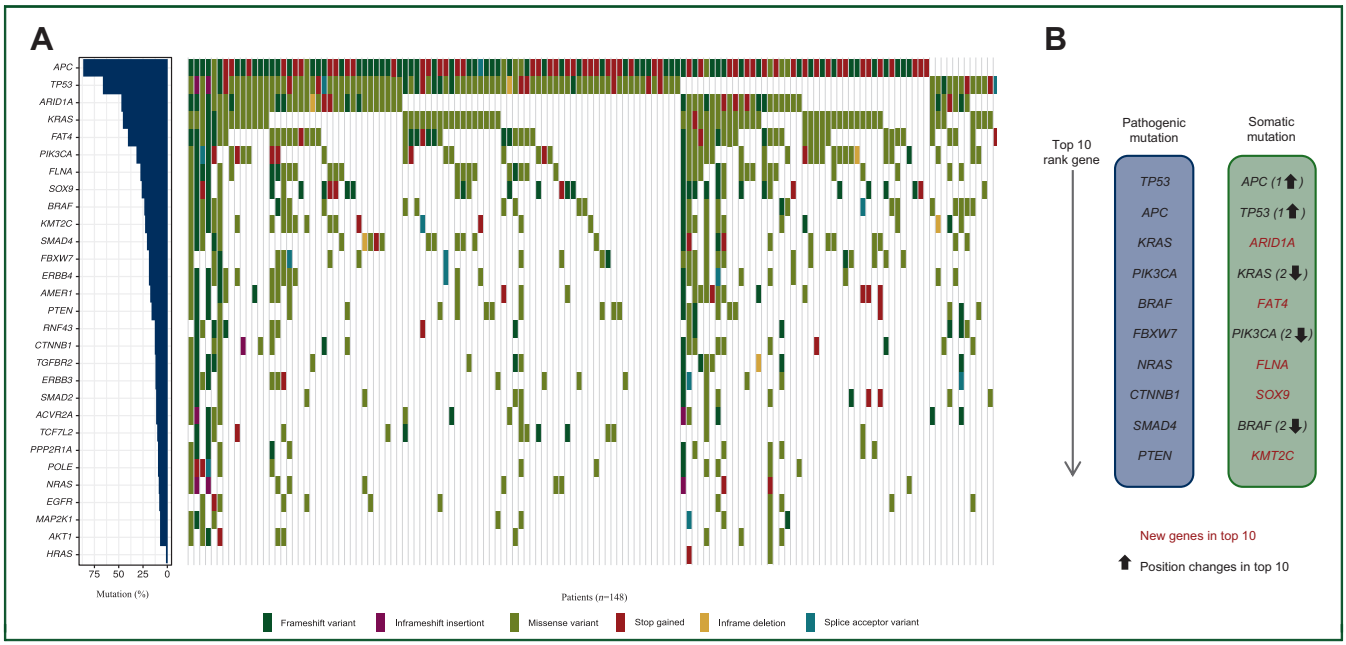

Figure 1. Somatic mutational landscape of 148 patients with localized colon cancer using the germline approach.
(A) Waterfall plot of somatic mutations detected in the cohort using a customized 29-gene panel in tumor samples. Synonymous and intronic variants are not included in the plot. (B) Top 10 mutated genes, segregated into two categories: pathogenic mutations[10] (left) and somatic mutations (right).

Droplet digital PCR analysis of cfDNA

Droplet digital PCR (ddPCR) was carried out on a QX200 ddPCR system (Bio-Rad) using TaqMan chemistry according to the instructions provided by the manufacturer. In summary, the mixture was prepared by combining up to 20 ng of cfDNA, 10 ml of a Supermix without dUTP (Bio-Rad), and 2 ml of FAM/HEX probes specific to the desired genetic changes. Emulsions were analyzed with the default settings for rare event detection and appropriate fluorophore detection. Probes for detecting specific alterations or simultaneous changes within a 25-pb range were designed using Bio-Rad platforms. A total of 95 ddPCR probes were required, with 59 probes used for pathogenic analysis and 75 probes for somatic analysis. All probes underwent validation using patient tissue samples.

Amplicon-based NGS for detection of ctDNA

For each patient under evaluation, we selected the four most common variants to design primer multiplexes. Library preparation followed a two-step PCR process consisting of 25 cycles for gene-specific PCR and 30 cycles for indexing PCR. The QIAGEN Multiplex PCR kit was used, and the libraries were purified using the MagSi-NGS PREP (magtivio) system. To ensure the primers were effective in detecting the chosen mutations, each multiplex was tested in the patient's tissue of origin. Libraries were quantified using the QuantiFluor dsDNA Assay Kit, and the size of the libraries was determined using the QIAxcel DNA Screening Gel (QIAGEN). Separate sequencing runs were conducted on a NextSeq instrument, using libraries from healthy controls as well as the corresponding cfDNA sample obtained at the postoperative timepoint, and using each primer multiplex. Patients showing higher measurements than healthy controls in at least two of the selected variants were classified as with MRD.

Statistical analysis

All analyses were carried out using R software version 4.3.0 (R Foundation, Vienna, Austria). Genomic variables and MRD are presented as frequencies and percentages. Differences in tumor-only versus germline testing were analyzed using the Wilcoxon signed-rank test. The association between quantitative variables was calculated by Pearson correlation. The median disease-free survival was estimated using the Kaplan—Meier method. Univariable analysis was carried out with Cox regression modeling. Plots were carried out using the ggplot2 and ComplexHeatmap packages in R. Shapiro—Wilk tests were carried out to evaluate normal distribution. P values were considered significant if $P < 0.05$ based on two-sided testing.

RESULTS

Genomic profiling of localized CC by sequencing primary tumor and matched WBCs

DNA extracted from tumor tissue and WBCs from a consecutive series of 148 patients with localized CC (stage I-III) was sequenced on a NextSeq platform at a median depth of 769× and 837×, respectively (Supplementary Materials, available at https://doi.org/10.1016/j.esmoop.2023.102051).[10] We examined somatic mutations in tumor tissue using an in-house designed panel of 29 frequently mutated genes in CC.[10] The mutations detected in tumor tissue were categorized as either germinal or somatic according to their presence in peripheral blood.

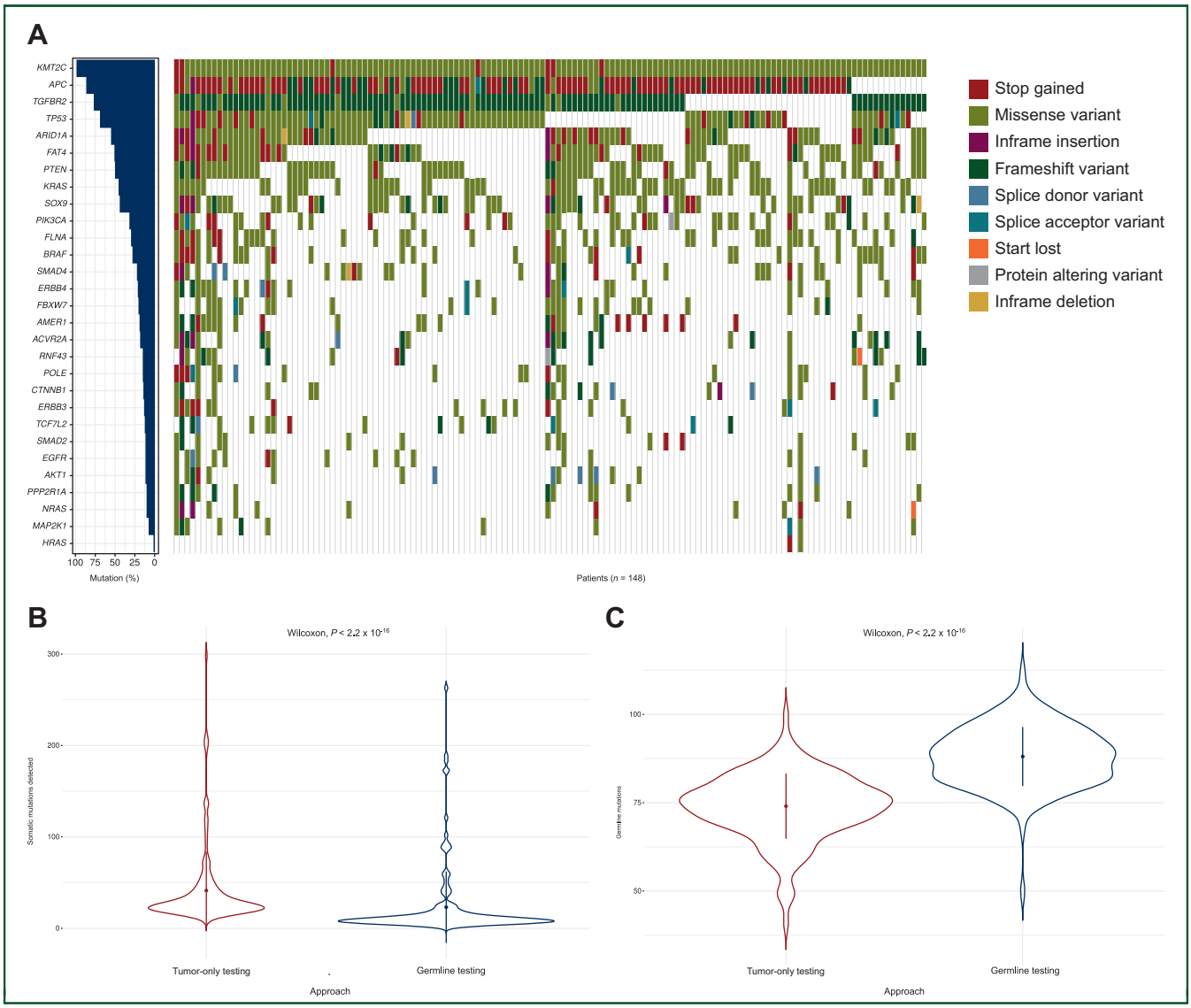

Figure 2. Somatic mutational landscape using the tumor-only testing versus germline testing approach.
(A) Waterfall plot of somatic mutations by the tumor-only approach classified by public databases. Synonymous and intronic variants are not included in the plot. (B) Violin plot comparing the number of somatic mutations obtained using the tumor-only approach (left) and the germline testing (right) approach. (C) Violin plot comparing the number of germline mutations obtained using the tumor-only approach (left) and the germline testing (right) approach.

We detected somatic mutations in all patients ($n = 148$; Figure 1A) using an optimized in-house bioinformatics pipeline to reduce background noise in variant calling. As expected, driver genes such as *APC* (86.45%), *TP53* (66.22%), *KRAS* (45.27%), *PIK3CA* (31.76%), and *BRAF* (23.65%) were detected among the top 10 mutated genes. In contrast to our previous analysis,[10] in which pathogenic mutations were only found in 132 patients (89.2%; median 2; range: 0-6), by adding WBC testing we could detect somatic mutations in all patients (median 11; range 1-234). This resulted in changes to the top 10 mutated genes with the highest allele mutation frequency (VAF; Figure 1B).

Added value of sequencing tumor tissue and paired WBCs

Among the major constraints in monitoring MRD are identifying molecular abnormalities to be followed in plasma at the time of diagnosis and the high frequency of false-positive DNA results.[8] Therefore we evaluated whether incorporating WBC sequencing data for classifying mutations as somatic or germinal variants would yield greater precision than relying on public genomic databases (which use only tumor data) and would also thus help lower the false-positive rate. The mutational landscape from both germline testing and tumor-only sequencing approaches is shown in Figures 1A and 2A, respectively. Our study cohort showed that the top 10 mutated genes varied depending on whether germline testing or tumor-only testing was analyzed. According to the tumor-only data, the most commonly mutated gene was *KMT2C* (146/148, 98.65%), whereas when germline testing was conducted, *KMT2C* was only present in 23% (34/148) of cases. A similar result was observed when analyzing *TGFBR2* mutations. It is noteworthy that well-established CC drivers such as *APC*

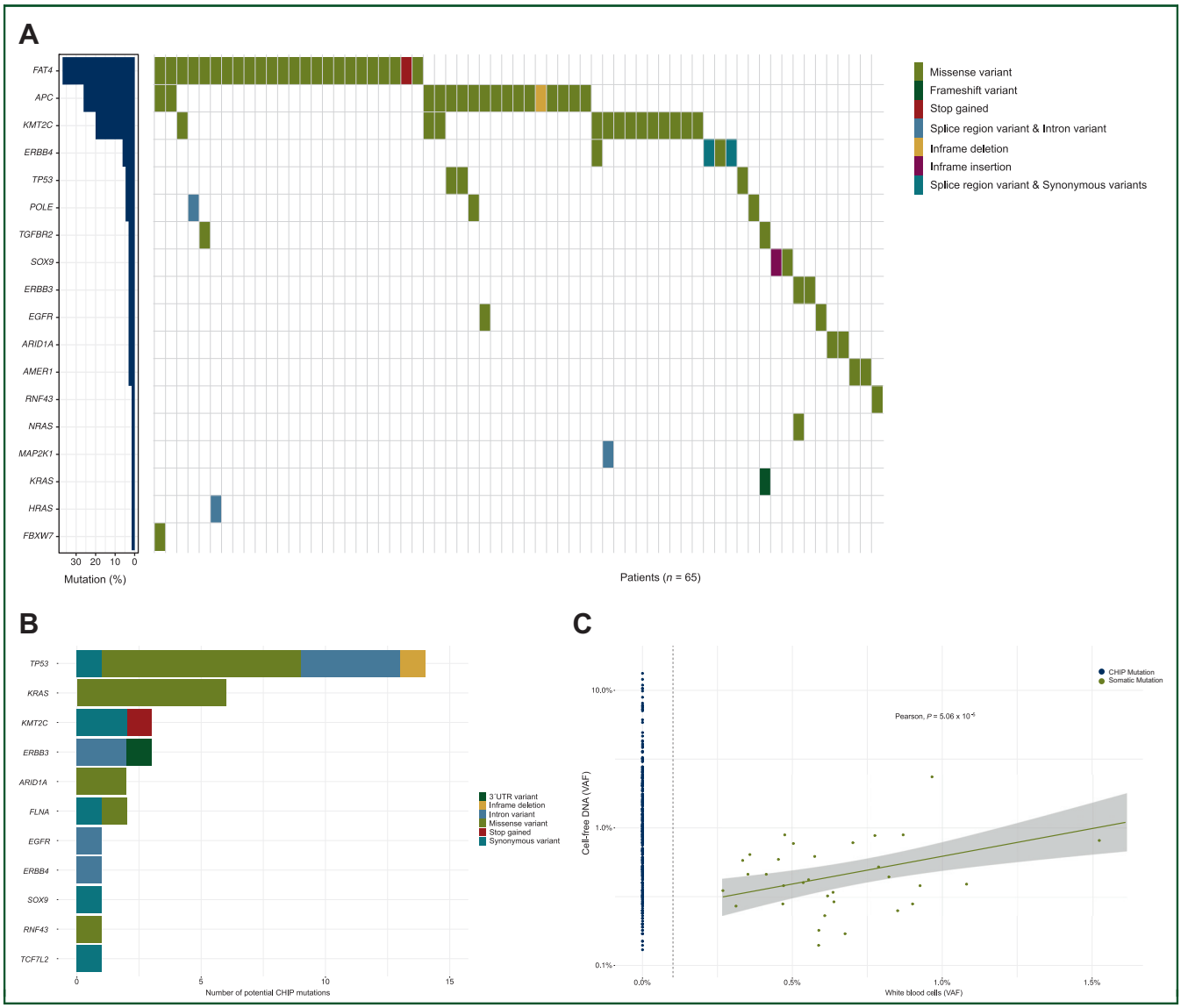

Figure 3. Mutations identified from white blood cell (WBC) sequencing.
(A) Waterfall plot of germline mutations detected in the cohort. (B) Number of clonal hematopoiesis of indeterminate potential (CHIP) mutations detected. (C) Correlation of variant allele frequency for CHIP mutations detected in cell-free DNA and WBCs.
UTR, untranslated region; VAF, variant allele frequency.

(86.49%), *TP53* (68.92%), *KRAS* (45.95%), and *PIK3CA* (31.76%) were less commonly found mutated when analyzing tumor-only data. Furthermore, a significantly higher number of somatic mutations were identified by analyzing tumor-only data compared with germline testing (Wilcoxon test, P value $= 2.2 \times 10^{-16}$; Figure 2B). As expected, the use of tumor-only data significantly decreased the germinal mutation count (Wilcoxon test, P value $= 2.2 \times 10^{-16}$; Figure 2C). These findings imply that carrying out germline sequencing could help prevent the possibility of misidentifying significant mutations, resulting in more accurate monitoring.[26]

Detection of germline alterations and clonal hematopoiesis

Although its primary purpose is to assist in treatment decision making, tumor testing can also uncover the existence of germline variants, which has important clinical implications for both patients and their relatives.[27] In our study, sequencing tumor tissue and paired WBCs enabled us to differentiate between somatic and germline mutations. Of 148 patients, nonsilent germline mutations were discovered in 65 individuals (43.92%; Figure 3A), with *FAT4* (11.49%), *APC* (9.46%), and *KMT2C* (6.08%) being the most commonly altered genes. We observed five distinct pathogenic variants in *APC*, with a prevalence rate ranging from 0.67% to 2.68%, and only one *TP53* variant, with a prevalence rate of 1.34%. In all cases, the VAF remained consistent in both WBC and tumor (Table 1). These findings are noteworthy for patients due to their potential association with hereditary cancer syndromes.[28]

Most *APC* alterations detected across our cohort (including p.Ser2621Cys, p.Ala2274Val, and Glu1317Gln) are listed as benign in public databases and are categorized as rare germline pathogenic variants.[29,30] Nevertheless,

Table 1. Pathogenic germline variants

Involved genes	Variant	Patients (%)	Median VAF in WBCs	Median VAF in tumor	Syndrome
APC	p.Ser2621Cys	1.34	0.49	0.50	Familial adenomatous polyposis
APC	p.Ser1400Leu	0.67	0.53	0.58	Attenuated familial adenomatous polyposis
APC	p.Ala2274Val	0.67	0.50	0.42	Gastric adenocarcinoma and proximal polyposis of the stomach
APC	p.Glu1317Gln	2.68	0.49	0.44	
APC	p.Ile1307Lys	1.34	0.49	0.56	
TP53	p.Val157Ile	1.34	0.48	0.61	Li–Fraumeni syndrome

Number of pathogenic germline mutations detected in the cohort. The number and percentage of patients affected by these mutations are also shown. The relationship of these mutations with hereditary cancers is shown on the right side of the table.
VAF, variant allele frequency; WBC, white blood cell.

variants such as p.Ser1400Leu and p.Ile1307Lys (observed in three of our patients) have been identified as leading factors for CC.[31-33] The p.Val157Ile *TP53* variant (found in two patients within the cohort) has been observed in patients with several tumor types, including sarcoma and breast cancer,[34,35] and has been linked to hereditary cancer in databases such as ClinVar.[36]

By contrast, germline testing reduced the false-positive rate of CHIP-associated mutations.[37] Our analysis of 118 plasma and matched WBC samples at the time of diagnosis revealed CHIP-related mutations in 27% of patients, with *TP53* (11.02%), *KRAS* (5.08%), and *KMT2C* (2.54%) exhibiting the highest VAF (Figure 3B). Consistent with previous studies,[38] the median mutant allele fraction among the 27 variants derived from WBCs was 0.62% (range 0.26%-1.83%), which was comparable to the median mutant allele fraction of the CHIP variants identified in cfDNA (median 0.44; range 0.14%-2.35%, $P = 0.07$, Wilcoxon's rank-sum test). The correlation between CHIP-related mutation levels in WBCs and in cfDNA (Pearson's correlation coefficient = 0.63) is shown in Figure 3C.

Identification of potentially actionable mutations

We analyzed the sequencing results of tumor tissue and paired WBCs which showed potentially actionable mutations, subsequently consulting OncoKB[39,40] to determine the level of evidence for a molecular-matched approach for each mutation detected (Table 2).[40] Of a total of 148 patients, 60 (40.54%) had one or more actionable targets, among which *KRAS* was the most frequently mutated gene (39.19%), followed by *PIK3CA* (21.62%). The percentages of *ARID1A* (9.46%), *NRAS* (6.08%), *BRAF* (5.41%), and *PTEN* (3.38%) alterations were smaller in the study population and compatible with previously published reports.[41] This highlights the significance of the *KRASG12C* mutation (2.7%) as both a driver and a target for novel compounds.

Table 2. Potentially actionable mutations detected in tumor samples

Genes	Druggable mutations (*n*)	Patients (%)	Treatment
PIK3CA	37	21.62	Alpelisib
ARID1A	18	9.46	PLX2853, Tazemetostat
NRAS	9	6.08	Binimetinib
BRAF	8	5.41	Selumetinib, Trametinib, Encorafenib, Dabrafenib, Vemurafenib
PTEN	6	3.38	GSK2636771, AZD8186
$KRAS^{G12C}$	4	2.70	Sotorasib, Adagrasib

Number of targetable mutations identified in the cohort. The table shows the treatment recommendations provided by the OncoKb database for each specific alteration.

ctDNA monitoring using tumor-only alterations compared with the addition of WBCs

After filtering alterations from WBCs, all individuals had at least one somatic mutation to be tracked in plasma, in contrast to only 89% of patients who underwent tumor-only sequencing (Supplementary Figure S1A, available at https://doi.org/10.1016/j.esmoop.2023.102051).

Postoperative ctDNA was analyzed in 68 patients using ddPCR based on patient-specific mutations to compare the monitoring performance between the tumor-only approach and the germline approach, in which the two most frequent somatic mutations were selected. Although the germline approach enabled the follow-up of more patients, there were no notable variations in the performance of the two techniques (proportion test, *P* value = 0.43; Supplementary Figure S1B, available at https://doi.org/10.1016/j.esmoop.2023.102051). To determine whether monitoring a greater number of mutations (beyond the two somatic mutations with the highest VAF) could enhance the accuracy of MRD detection, all somatic mutations with an available ddPCR probe were assessed. This latter approach improved accuracy, but again no significant differences in postoperative MRD detection were observed (proportion test, *P* value = 0.56).

NGS-based amplicons were also analyzed to enhance the sensitivity of MRD detection.[42,43] However, no significant differences were observed compared with the ddPCR technique (proportion test, $P = 1$; Supplementary Figure S1B, available at https://doi.org/10.1016/j.esmoop.2023.102051).

NGS and ddPCR analysis were in good agreement in postoperative ctDNA in the mutational profile (Supplementary Figure S3A, available at https://doi.org/10.1016/j.esmoop.2023.102051) and the VAF ($P < 0.0001$, Supplementary Figure S3B, available at https://doi.org/10.1016/j.esmoop.2023.102051). MRD detection in postoperative plasma using ddPCR showed an HR of 4.59 ($P = 0.0005$; Supplementary Figure S4A, available at https://doi.org/10.1016/j.esmoop.2023.102051), whereas the amplicon-based NGS technique had an HR of 5.12 ($P =$

0.0037; Supplementary Figure S4B, available at https://doi.org/10.1016/j.esmoop.2023.102051) for disease-free survival.

DISCUSSION

Colorectal cancer is a significant public health problem and the second leading cause of cancer-related deaths worldwide.[1] The most dramatic statistics concern the younger population; the incidence of this cancer is estimated to double among under 50-year old by 2030.[44] This highlights a need to more precisely identify patients at a higher risk of relapse and to develop better adjuvant treatment. Pathologic staging has so far served to help clinicians identify patients at risk of recurrence.[2] However, ∼15%-30% of patients experience relapse despite receiving optimal initial treatment.

ctDNA has emerged as a promising prognostic and possibly predictive biomarker in the management of patients with localized CRC. ctDNA has shown various potential applications, including MRD detection, early recurrence monitoring, molecular profiling, and prediction of therapeutic response.[45] Nonetheless, liquid biopsy has certain limitations, such as false-positive and false-negative results in ctDNA, which pose challenges in detecting MRD.[3-12]

The aims of this study were to reduce the false-positive rate in MRD detection and to identify new variants through paired tumor tissue and WBC sequencing. In our previous study, in which only tissue from the primary tumor was sequenced, only 89% of patients had at least one mutation that could be tracked in plasma during follow-up.[10] Because of the absence of WBCs to ensure a reliable NGS result, only pathogenic variants were selected. As a result, 11% of patients lacked mutations that could be monitored in plasma, thus precluding MRD evaluation. By contrast, matched tumor tissue and WBCs sequencing (hereafter referred to as germline testing) not only permits the detection of any somatic mutations, whether pathogenic or otherwise, but can also distinguish them from germline mutations. This approach enabled the detection of at least one somatic mutation for monitoring in plasma in all (100%) patients in our sample, thereby increasing the number of patients eligible for MRD tracking.

In concurrence with the literature, the top 10 most frequently mutated genes occur in driver genes in CC such as *TP53*, *APC*, *KRAS, PIK3CA,* and *BRAF*, regardless of whether only pathogenic variants or any somatic mutations are considered.[39] Taking the latter data into account, however, resulted in changes to the top 10 mutated genes, with the emergence of five new genes (*ARID1A*, *FAT4*, *FLNA*, *SOX9*, and *KMT2C*) not initially identified. This points to the importance of germline testing to detect further somatic mutations.

It should be noted that in this study, we used a custom panel of 29 genes frequently mutated in CRC,[10] acknowledging the limitations inherent to this method. To assess the theoretical detection sensitivity of the sequencing approach in CC, we determined the proportion of CC in The Cancer Genome Atlas database[46] with alterations in one or more of the 29 analyzed genes. These analyses showed that our targeted panel would have a sensitivity of ∼82.4%, given that 435 of 528 CC cases had at least one alteration in these genes (Supplementary Figure S5, available at https://doi.org/10.1016/j.esmoop.2023.102051). In addition, we detected at least one mutation at diagnosis in all patients in our cohort, which contrasts with other tumors such as breast cancer where the lower molecular alteration detection rate reduces the number of patients eligible to be monitored in plasma and thus the MRD detection rate.[47] In addition, the sensitivity to detect ctDNA immediately after surgery using our CC driver gene panel is 53%, slightly lower than has been reported in commercial assays.[11] This sensitivity increases throughout mutation tracking due to the increase in tumor burden over time.[10]

Recently, the recommendations of two groups of experts in the field of ESMO and ASCO have pointed to next-generation sequencing-based multigene assays as the preferred method for detecting MRD, as opposed to ddPCR-based assays.[42,43] For this reason, we compared ddPCR and amplicon-based NGS for sensitivity to MRD. Unfortunately, no improvement in MRD accuracy was observed using amplicon-based NGS, probably due to the small number of patients analyzed.[48]

The prevalence of CHIP mutations increases as individuals age, and these mutations have been found at low levels in up to 95% of noncancer patients aged 50-60 years.[49] Initial studies have focused primarily on CHIP mutations in genes known to be associated with hematological malignancies, such as *TET2*, *DNMT3*, *JAK2*, and *ASXL1,* which have limited relevance in solid tumors such as CC. However, subsequent studies have also identified mutations in other genes, including *TP53* and *KRAS*, which could contribute to false-positive results when detecting MRD in patients with CC.[50] To decrease the false-positive rate, we sequenced cfDNA paired with WBCs and identified mutations in CHIP-associated genes in 27% of our cohort, with *TP53*, *KRAS*, and *KMT2C* being the most frequently mutated genes according to the literature.[51] Use of WBCs therefore ensures precise interpretation of genetic variations to accurately detect MRD.

Germline sequencing may also have certain ethical implications.[27] In our cohort, 9.2% of individuals presented pathogenic or likely pathogenic mutations in genes such as *TP53* or *APC*, which are associated with hereditary cancer syndromes such as Li—Fraumeni, familial adenomatous polyposis, attenuated familial adenomatous polyposis and gastric adenocarcinoma, and proximal polyposis of the stomach.[52]

Likewise, using germline sequencing we could identify actionable mutations that could potentially improve MRD management, coinciding with interventional targeted therapy trials in the advanced setting.[53] We observed that 40% of the cohort had targetable alterations at diagnosis, including oncogenic mutations in the RAS/RAF, PI3K, and WNT pathways and tumor suppressor pathways. Patients with actionable molecular aberrations may derive greater

benefit from targeted therapies tailored to their specific alterations than from standard treatments. Accurate selection of targeted agents based on molecular alterations can significantly enhance treatment decisions and overall management, particularly in patients with advanced cancer.[41]

The assay used in this study is based on a tumor-informed approach. The two techniques currently used for MRD assessment are the tumor-informed approach and the tumor-agnostic approach. To date, no comparative data are available on these two assays in terms of MRD detection sensitivity, but ongoing prospective trials will provide more information in this regard.[54,55]

CONCLUSIONS

Our findings demonstrate that utilizing both tumor DNA and paired WBCs for sequencing not only expands the pool of patients eligible for monitoring MRD in plasma, but also reduces the frequency of false-positive results.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of the INCLIVA Biobank (PT20/00029; B.0000768 ISCIII) as part of the Valencian Biobanking Network and the Spanish National Biobanks Network, the Precision Medicine Unit of the INCLIVA Biomedical Research Institute, and especially all the patients and their families for their participation in this study.

FUNDING

This work was supported by the Instituto de Salud Carlos III [grant number PI21/00689], the Spanish Society of Medical Oncology (SEOM) [no grant number], The Spanish Cooperative Group for Digestive Tumor Therapy (TTD Group), the Spanish Association Against Cancer (AECC) [grant number GCAEC20030CERV], and Mutua Madrileña Foundation. Joan Rodés contract [grant number JR20/00005 to NT and grant number JR21/00042 to VG]. A Generalitat Valenciana fellowship program [grant number APOSTD/2021/168 to FGV]. Generalitat Valenciana [grant number ACIF2020/381 to JMA]. A PFIS contract from ISCIII [grant number FI22/00267 to BGM]. Ministerio de Universidades [grant number FPU21/00373 to BMC]. Part of the equipment used in this study has been funded by Generalitat Valenciana and co-financed with ERDF funds [grant number OP ERDF of Comunitat Valenciana 2014-2020]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

DISCLOSURE

AC declares institutional research funding from Genentech, Merck Serono, Bristol Myers Squibb, MSD, Roche, BeiGene, Bayer, Servier, Lilly, Natera, Novartis, Takeda, Astellas, and Fibrogen; and advisory board or speaker fees from Merck Serono, Roche, Servier, Takeda, and Astellas. NT declares advisory board or speaker fees from Merck Serono, Servier, Pfizer, Natera, and Guardant Health. VG declares advisory board fees from Boehringer Ingelheim and institutional research funding from Bayer, Boehringer, Roche, Genentech, Merck Serono, BeiGene, Servier, Lilly, Novartis, Takeda, Astellas, Fibrogen, amcure, Natera, Sierra Oncology, Astra Zeneca, Medimmune, BMS, and MSD. SR declares personal fees as an invited speaker from Amgen, MSD, and Servier; advisory board fees from Amgen, Servier, and Sirtex; and institutional funding from Ability Pharmaceuticals, Astellas, G1 Therapeutics, Hutchinson, Menarini, Mirati, Novartis, Pfizer, Pierre Fabre, Roche, and Seagen. All other authors have declared no conflicts of interest.

DATA SHARING

All sequencing data supporting the findings in this study are currently being submitted to the European Genome-phenome Archive (EGA). The accession number will be provided once the deposition process is completed.

CODE AVAILABILITY

The code for analyses is available at https://github.com/jmartinarana/INCLIVA-CC-panel. The datasets used and/or analyzed to generate the figures in the current study are also accessible in the repository.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209-249.
2. Argilés G, Tabernero J, Labianca R, et al. Localised colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2020;31(10):1291-1305.
3. Tie J, Wang Y, Tomasetti C, et al. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. *Sci Transl Med*. 2016;8(346):346ra92.
4. Reinert T, Henriksen TV, Christensen E, et al. Analysis of plasma cell-free DNA by ultradeep sequencing in patients with stages I to III colorectal cancer. *JAMA Oncol*. 2019;5(8):1124-1131.
5. Tie J, Cohen JD, Wang Y, et al. Circulating tumor DNA analyses as markers of recurrence risk and benefit of adjuvant therapy for stage III colon cancer. *JAMA Oncol*. 2019;5(12):1710-1717.
6. Tarazona N, Gimeno-Valiente F, Gambardella V, et al. Targeted next-generation sequencing of circulating-tumor DNA for tracking minimal residual disease in localized colon cancer. *Ann Oncol*. 2019;30(11):1804-1812.
7. Taieb J, Taly V, Henriques J, et al. Prognostic value and relation with adjuvant treatment duration of ctDNA in stage III colon cancer: a *post hoc* analysis of the PRODIGE-GERCOR IDEA-France trial. *Clin Cancer Res*. 2021;27(20):5638-5646.
8. Henriksen TV, Tarazona N, Frydendahl A, et al. Circulating tumor DNA in stage III colorectal cancer, beyond minimal residual disease detection, toward assessment of adjuvant therapy efficacy and clinical behavior of recurrences. *Clin Cancer Res*. 2022;28(3):507-517.
9. Taniguchi H, Nakamura Y, Kotani D, et al. CIRCULATE-Japan: circulating tumor DNA-guided adaptive platform trials to refine adjuvant therapy for colorectal cancer. *Cancer Sci*. 2021;112(7):2915-2920.
10. Tarazona N, Gimeno-Valiente F, Gambardella V, et al. Detection of postoperative plasma circulating tumour DNA and lack of CDX2 expression as markers of recurrence in patients with localised colon cancer. *ESMO Open*. 2020;5(5):e000847.
11. Kotani D, Oki E, Nakamura Y, et al. Molecular residual disease and efficacy of adjuvant chemotherapy in patients with colorectal cancer. *Nat Med*. 2023;29(1):127-134.

12. Mo S, Ye L, Wang D, et al. Early detection of molecular residual disease and risk stratification for stage I to III colorectal cancer via circulating tumor DNA methylation. *JAMA Oncol*. 2023;9(6):770-778.
13. Tie J, Cohen JD, Lahouel K, et al. Circulating tumor DNA analysis guiding adjuvant therapy in stage II colon cancer. *N Engl J Med*. 2022;386(24):2261-2272.
14. van der Leest P, Schuuring E. The potential of combined mutation sequencing of plasma circulating cell-free DNA and matched white blood cells for treatment response prediction. *Mol Oncol*. 2020;14(3):487-489.
15. Newman AM, Lovejoy AF, Klass DM, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat Biotechnol*. 2016;34(5):547-555.
16. Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med*. 2014;6(224):224ra24.
17. Andrews S. FastQC: a quality control tool for high throughput sequence data. *Babraham Bioinform*. 2010. Available at https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed October 20, 2023.
18. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 2011;17.
19. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. 2011;27(6):863-864.
20. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760.
21. FGBIO's Best Practice Consensus Pipeline, version 1.1.0. Available at https://github.com/fulcrumgenomics/fgbio/. Accessed October 20, 2023.
22. McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303.
23. Xu C, Gu X, Padmanabhan R, et al. smCounter2: an accurate low frequency variant caller for targeted sequencing data with unique molecular identifiers. *Bioinformatics*. 2019;35(8):1299-1309.
24. Wilm A, Aw PP, Bertrand D, et al. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res*. 2012;40(22):11189-11201.
25. McLaren W, Gil L, Hunt SE, et al. The ensembl variant effect predictor. *Genome Biol*. 2016;17(1):122.
26. Berchuck JE, Boiarsky D, Silver R, et al. Addition of germline testing to tumor-only sequencing improves detection of pathogenic germline variants in men with advanced prostate cancer. *JCO Precis Oncol*. 2022;6:e2200329.
27. Borno HT, Rider JR, Gunn CM. The ethics of delivering precision medicine-pretest counseling and somatic genomic testing. *JAMA Oncol*. 2020;6(6):815-816.
28. Stenson PD, Ball EV, Mort M, et al. Human gene mutation database (HGMD): 2003 update. *Hum Mutat*. 2003;21(6):577-581.
29. Bodian DL, McCutcheon JN, Kothiyal P, et al. Germline variation in cancer-susceptibility genes in a healthy, ancestrally diverse cohort: implications for individual genome sequencing. *PLoS One*. 2014;9(4):e94554.
30. Minde DP, Anvarian Z, Rüdiger SG, Maurice MM. Messing up disorder: how do missense mutations in the tumor suppressor protein APC lead to cancer? *Mol Cancer*. 2011;10:101.
31. White S, Bubb VJ, Wyllie AH. Germline APC mutation (Gln1317) in a cancer-prone family that does not result in familial adenomatous polyposis. *Genes Chromosomes Cancer*. 1996;15(2):122-128.
32. Chang YS, Lee CC, Ke TW, et al. Molecular characterization of colorectal cancer using whole-exome sequencing in a Taiwanese population. *Cancer Med*. 2019;8(8):3738-3747.
33. Gryfe R, Di Nicola N, Gallinger S, Redston M. Somatic instability of the APC I1307K allele in colorectal neoplasia. *Cancer Res*. 1998;58(18):4040-4043.
34. Mitchell G, Ballinger ML, Wong S, et al. High frequency of germline TP53 mutations in a prospective adult-onset sarcoma cohort. *PLoS One*. 2013;8(7):e69026.
35. Jalkh N, Chouery E, Haidar Z, et al. Next-generation sequencing in familial breast cancer patients from Lebanon. *BMC Med Genomics*. 2017;10(1):8.
36. Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*. 2016;44(D1):D862-D868.
37. Leal A, van Grieken NCT, Palsgrove DN, et al. White blood cell and cell-free DNA analyses for detection of residual disease in gastric cancer. *Nat Commun*. 2020;11(1):525.
38. Li BT, Janku F, Jung B, et al. Ultra-deep next-generation sequencing of plasma cell-free DNA in patients with advanced lung cancers: results from the Actionable Genome Consortium. *Ann Oncol*. 2019;30(4):597-603.
39. Chakravarty D, Gao J, Phillips SM, et al. OncoKB: a precision oncology knowledge base. *JCO Precis Oncol*. 2017;2017:PO.17.00011.
40. Mateo J, Chakravarty D, Dienstmann R, et al. A framework to rank genomic alterations as targets for cancer precision medicine: the ESMO Scale for Clinical Actionability of molecular Targets (ESCAT). *Ann Oncol*. 2018;29(9):1895-1902.
41. Gambardella V, Lombardi P, Carbonell-Asins JA, et al. Molecular profiling of advanced solid tumours. The impact of experimental molecular-matched therapies on cancer patient outcomes in early-phase trials: the MAST study. *Br J Cancer*. 2021;125(9):1261-1269.
42. Dasari A, Morris VK, Allegra CJ, et al. ctDNA applications and integration in colorectal cancer: an NCI Colon and Rectal-Anal Task Forces whitepaper. *Nat Rev Clin Oncol*. 2020;17(12):757-770.
43. Pascual J, Attard G, Bidard FC, et al. ESMO recommendations on the use of circulating tumour DNA assays for patients with cancer: a report from the ESMO Precision Medicine Working Group. *Ann Oncol*. 2022;33(8):750-768.
44. Mauri G, Sartore-Bianchi A, Russo AG, Marsoni S, Bardelli A, Siena S. Early-onset colorectal cancer in young individuals. *Mol Oncol*. 2019;13(2):109-131.
45. Malla M, Loree JM, Kasi PM, Parikh AR. Using circulating tumor DNA in colorectal cancer: current and evolving practices. *J Clin Oncol*. 2022;40(24):2846-2857.
46. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;487(7407):330-337.
47. Garcia-Murillas I, Chopra N, Comino-Méndez I, et al. Assessment of molecular relapse detection in early-stage breast cancer. *JAMA Oncol*. 2019;5(10):1473-1478.
48. Roberto TM, Jorge MA, Francisco GV, Noelia T, Pilar RG, Andrés C. Strategies for improving detection of circulating tumor DNA using next generation sequencing. *Cancer Treat Rev*. 2023;119:102595.
49. Young AL, Challen GA, Birmann BM, Druley TE. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. *Nat Commun*. 2016;7:12484.
50. Gibson CJ, Steensma DP. New insights from studies of clonal hematopoiesis. *Clin Cancer Res*. 2018;24(19):4633-4642.
51. Severson EA, Riedlinger GM, Connelly CF, et al. Detection of clonal hematopoiesis of indeterminate potential in clinical sequencing of solid tumor specimens. *Blood*. 2018;131(22):2501-2505.
52. Garutti M, Foffano L, Mazzeo R, et al. Hereditary cancer syndromes: a comprehensive review with a visual tool. *Genes (Basel)*. 2023;14(5):1025.
53. Nakamura Y, Taniguchi H, Ikeda M, et al. Clinical utility of circulating tumor DNA sequencing in advanced gastrointestinal cancer: SCRUM-Japan GI-SCREEN and GOZILA studies. *Nat Med*. 2020;26(12):1859-1864.
54. Tarazona N, Gimeno-Valiente F, Cervantes A. Minimal residual disease in gastroesophageal adenocarcinoma: the search for the invisible. *ESMO Open*. 2022;7(4):100547.
55. Kasi PM. Tumor-informed versus plasma-only liquid biopsy assay in a patient with multiple primary malignancies. *JCO Precis Oncol*. 2022;6:e2100298.