Molecular and functional profiling unravels targetable vulnerabilities in colorectal cancer

Efstathios-Iason Vlachavas[1] (iD), Konstantinos Voutetakis[2] (iD), Vivian Kosmidou[2],
Spyridon Tsikalakis[1], Spyridon Roditis[3,4], Konstantinos Pateas[3], Ryangguk Kim[5], Kymberleigh Pagel[6],
Stephan Wolf[7], Gregor Warsow[8], Antonia Dimitrakopoulou-Strauss[9], Georgios N. Zografos[3],
Alexander Pintzas[2], Johannes Betge[10,11,12], Olga Papadodima[2] (iD) and Stefan Wiemann[1] (iD)

1 Division of Molecular Genome Analysis, German Cancer Research Center, Heidelberg, Germany
2 Institute of Chemical Biology, National Hellenic Research Foundation, Athens, Greece
3 3rd Surgical Department G.Gennimatas Hospital, Athens, Greece
4 Surgical Department, University Hospital of North Midlands, Stoke-on-Trent, UK
5 Oak Bioinformatics, LLC, Fairfax, VA, USA
6 SpinSys-Diné, Falls Church, VA, USA
7 High-Throughput Sequencing Core Facility, German Cancer Research Center, Heidelberg, Germany
8 Omics IT and Data Management Core Facility, German Cancer Research Center, Heidelberg, Germany
9 Clinical Cooperation Unit Nuclear Medicine, German Cancer Research Center, Heidelberg, Germany
10 Junior Clinical Cooperation Unit Translational Gastrointestinal Oncology and Preclinical Models, German Cancer Research Center (DKFZ), Heidelberg, Germany
11 Department of Medicine II, University Medical Center Mannheim, Medical Faculty Mannheim, Heidelberg University, Germany
12 DKFZ-Hector Cancer Institute at University Medical Center Mannheim, Germany

Colorectal cancer (CRC) patients with microsatellite-stable (MSS) tumors are mostly treated with chemotherapy. Clinical benefits of targeted therapies depend on mutational states and tumor location. Many tumors carry mutations in KRAS proto-oncogene, GTPase (*KRAS*) or B-Raf proto-oncogene, serine/threonine kinase (*BRAF*), rendering them more resistant to therapies. We performed whole-exome sequencing and RNA-Sequencing of 28 tumors of the Athens Comprehensive Cancer Center CRC cohort, and molecularly characterized CRC patients based on their microsatellite instability (MSI) status, single-nucleotide variations (SNVs)/copy number alterations (CNAs), and pathway/transcription factor activities at the individual patient level. Variants were classified using a computational score for integrative cancer variant annotation and prioritization. Complementing this with public multi-omics datasets, we identified activation of transforming growth factor beta (TGFβ) signaling to be more strongly activated in MSS patients, whereas Janus kinase (JAK)–signal transducer and activator of transcription (STAT) and mitogen-activated protein kinase (MAPK) molecular cascades were activated specifically in MSI tumors. We

Abbreviations

ACCC, Athens Comprehensive Cancer Center; CMS, consensus molecular subtypes; CNAs, copy number alterations; COAD, colorectal adenocarcinoma dataset of the TCGA; COSMIC, catalog of somatic mutations in cancer database; CPTAC, Clinical Proteomic Tumor Analysis Consortium; CRC, colorectal cancer; DKFZ-ODCF, German Cancer Research Center-Omics IT and Data Management Core Facility; EMT, epithelial-to-mesenchymal transition; HCPC, hierarchical clustering on principal components; InDels, insertions and deletions; IntOGen, Integrative OncoGenomics database; MANE, matched annotation from NCBI and EMBL-EBI; MASTER, molecularly aided stratification for tumor eradication research; MSI, microsatellite instable; MSS, microsatellite stable; NCT/DKTK, National Center for Tumor Diseases/German Cancer Consortium; PROGENy, pathway responsive genes R package; PTM-SEA, post-translational modification set enrichment analysis; READ, rectum adenocarcinoma cohort of the TCGA; RNAseq, RNA sequencing; RTKs, Receptor Tyrosine Kinases; SBS, COSMIC – mutational signatures; SNVs, single nucleotide variations; ssGSEA, single sample gene set enrichment analysis; TCGA, The Cancer Genome Atlas; TFs, transcription factors; TMB, tumor mutational burden; WES, whole exome sequencing; WT, wild-type.

Im Neuenheimer Feld 580, 69120
Heidelberg, Germany
E-mail: s.wiemann@dkfz-heidelberg.de

Efstathios-Iason Vlachavas and Konstantinos
Voutetakis contributed equally to this article.

unraveled mechanisms consistently perturbed in the transcriptional and mutational circuits and identified Runt-related transcription factors (RUNX transcription factors) as putative biomarkers in CRC, given their role in the regulation of pathways involved in tumor progression and immune evasion. Assessing the immunogenicity of CRC tumors in the context of RAS/RAF mutations and MSI/MSS status revealed a critical impact that *KRAS* mutations have on immunogenicity, particularly in the MSS patient subgroup, with implications for diagnosis and treatment.

1. Introduction

Colorectal cancer (CRC) is a highly heterogeneous disease associated with different molecular background and clinical management. Improved screening strategies, surgical techniques and advanced treatment protocols have led to better clinical outcomes of CRC patients in the last decades [1,2]. The recent development of covalent inhibitors specifically blocking the KRAS G12C variant have indicated significant progress in targeting even previously 'undruggable' onco-proteins [3]. Despite these advances, CRC has remained the second and third leading cause of cancer-related death in the US and worldwide, respectively [4,5]. Furthermore, recent epidemiological data has indicated a concerning trend of increased CRC incidence particularly in younger individuals below the age of 55, coupled with more advanced disease stages [5]. The response to treatment is highly variable and not durable in the majority of cases. This dismal clinical perspective is impacted by various factors, including tissue tropism effects that allele-specific *RAS* mutants have in combination with other actionable alterations [6], clonal heterogeneity [7], phenotypic plasticity [8], location of tumor [9], and the tumor microenvironment [10]. Collectively, these findings highlight the need for novel therapies focused specifically toward the personalized treatment of CRC. High-throughput transcriptomic profiling and large-scale genome sequencing have led to substantial progress in molecular characterization of the disease. In addition, alterations in a number of genes with critical roles in CRC have been identified, such as in *APC*, *BRAF*, *EGFR*, *NRAS*, *KRAS*, and *PIK3CA* [11,12]. The resulting aberrant activation of signaling pathways is increasingly exploited in precision therapies. Recent accomplishments include combinatorial strategies targeting co-activated pathways in various types of unresectable CRC such as BRAF V600E mutated [2] and ERBB2-positive carcinomas [1].

The microsatellite instability (MSI) status is among the most informative parameters determining treatment decision toward personalized medicine in CRC [13]. According to the MSI status, CRC is categorized into three types: high microsatellite instable (MSI-H), low microsatellite instable (MSI-L) and microsatellite stable (MSS) [13]. MSI-L and MSS tumors are commonly grouped into one type [14]. Accordingly, we use MSS for MSI-L/MSS and MSI for MSI-H tumors throughout. MSI tumors account for ~ 15% of CRC cases and are associated with sporadic or hereditary disease etiologies. While the former is associated with *BRAF* mutations [15] and *MLH1* promoter methylation [16], the latter has been associated with mutations mostly in mismatch repair genes *MLH1*, *MSH2*, *MSH6* as well as *PMS2* and is classified as Lynch syndrome [17]. MSI tumors have a high mutational burden, resulting in the expression of neo-antigens that may be recognized by the immune system. Differences in the immune microenvironment thus distinguish MSI and MSS CRCs, and have immediate therapeutic impact: MSI CRC tumors mostly respond to immunotherapy, which has revolutionized clinical practice and outcome of respective patients [18].

In contrast, MSS tumors, which account for the majority of colorectal cancers, generally display chromosomal rather than microsatellite instability. These tumors are often characterized by mutations in key tumor suppressor and oncogene genes such as *APC* as well as *TP53* and *KRAS*, respectively. These tumors tend to have a lower tumor mutational burden (TMB) and reduced immune infiltration compared to MSI tumors. MSS CRC patients commonly do not benefit from immunotherapies and have mostly limited targeted treatment options, wherein the most informative molecular distinction involves the RAS mutation status [19]. Despite extensive research conducted on the molecular differences between MSS and MSI

colorectal tumors [20], the therapeutic management of MSS, along with the heterogeneity often observed within both MSS and MSI tumors [21], underscores the urgent need for the integration of different molecular layers and a focus on personalized profiling.

In this study, we aimed to shed light on the multi-layered differences in MSI/MSS tumors, their downstream signaling effects, and their interrelation with *RAS/RAF* mutations in regard to immunogenicity and pathway activation [20]. Hence, by leveraging a prospective patient cohort of the Athens Comprehensive Cancer Center (ACCC—https://www.accc.gr/index.html; accessed: 2024/01/02) and two public cancer genomics datasets, we aspired to elucidate the molecular complexity of microsatellite (in)stability status in CRC and to uncover biological mechanisms that might be exploited to enhance therapeutic options, particularly for patients with the MSS type.

2. Materials and methods

2.1. Patients and clinical specimens

Colorectal cancer patients operated in the surgical department of Gennimatas General Hospital in Athens, which is part of the Athens Comprehensive Cancer Center (ACCC), participated in the study. The study received approval from the Ethics Committee of Gennimatas Hospital (protocol number 27659) and was conducted in accordance with the principles outlined in the Declaration of Helsinki. Written informed consent was obtained from all patients. Only patients ($n = 28$) with histologically confirmed primary colorectal adenocarcinoma and one case of rectal cancer who had received radiotherapy before surgery were enrolled in the study cohort (between August 2018 and October 2020). Specimens from each patient were obtained from both the primary tumor and corresponding normal mucosa, which was taken from the histologically negative margins (at least 5 cm away from the macroscopically defined cancerous area) to ensure the absence of cancer cells. Upon surgery, resected samples were placed in 5 mL of RNA*later* (Ambion, Austin, TX, USA) to preserve RNA integrity.

2.2. DNA and RNA extraction

DNA and RNA were isolated with the AllPrep DNA/RNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Quantification of nucleic acids was performed with a Qubit 2.0 Fluorometer (ThermoFisher Scientific, Waltham, MA, USA) using the Qubit 1× dsDNA BR Assay (ThermoFisher Scientific) and RNA BR Assay kits (ThermoFisher Scientific) for DNA and RNA, respectively. In total, 28 paired tumor and normal DNA samples were isolated as well as 28 matching tumor RNA samples.

2.3. Whole exome sequencing (WES) and variant calling

DNA libraries were prepared from tumor and matched normal tissues of ACCC patients using Low Input Exome-Seq Human v6 (Agilent Technologies, Waldbronn, Germany) for exome capture and analyzed on a NovaSeq 6000 (Paired-End 150 bp S2, Illumina, San Diego, CA, USA) in the DKFZ High-Throughput Sequencing Core Facility. The resulting raw sequencing reads (median: 123 million reads per sample; spread: 80–148 million reads per sample) were preprocessed based on the internal WES alignment workflow (AlignmentAndQCWorkflows: 1.2.73-1), as part of the DKFZ-ODCF OTP pipeline [22]. Briefly, after initial quality control using FastQC, raw reads were aligned to a hg19 customized version of the 1000 genomes reference genome (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/phase2_reference_assembly_sequence/hs37d5.fa.gz) (1KGRef_PhiX), using BWA-MEM (version 0.7.15). PhiX-genome sequence was added to the reference as an additional contig to remove spike-in sequences used to identify potential cross-contaminations and sample swaps. Marking of duplicated reads was performed with SAMBAMBA tools (https://github.com/biod/sambamba/releases, version 0.6.5) embedded in the DKFZ-ODCF Roddy pipeline (https://github.com/DKFZ-ODCF/AlignmentAndQCWorkflows).

High-confidence functional somatic single nucleotide variants (SNVs) and InDels (Insertions & Deletions) were identified using DKFZ-ODCF SNVCallingWorkflow (https://github.com/DKFZ-ODCF/SNVCallingWorkflow; version 1.2.166-1) and IndelCallingWorkflow (https://github.com/DKFZ-ODCF/IndelCallingWorkflow; version 1.2.177-0), respectively, based on the BAM files of each tumor and its paired control sample. Various quality control criteria were performed to exclude spurious or/and low-quality calls (such as identification of low-quality bases, variant allele fraction and presence of strand bias artifacts), finally assigning a confidence value to each variant. Sample ACCC_CRC_02 was disregarded in further analysis as COSMIC SBS signature [23] analysis (v3.4-October 2023) flagged this as having a potential PCR/sequencing artifact-based strand bias (prevalence of C>A alterations indicated presence of COSMIC SBS52; https://cancer.sanger.ac.uk/signatures/sbs/sbs52/; accessed 2023/12/21). Finally,

copy numbers. Allowed means that balanced segments were restricted to even total copy number states whereas any integer copy number state was allowed for imbalanced segments and allele-specific copy numbers. TCC/ploidy combinations requiring negative copy number states or DH values larger than 1 for any segment were excluded. Local minima in the weighted mean distance were considered as possible TCC/ploidy solutions for the sample and were visually evaluated.

2.6. RNA-sequencing and analysis

RNAseq libraries were prepared using the TruSeq Stranded library protocol (Illumina) and sequencing was done with a NovaSeq 6000 (Paired-end 100 bp S1) in the DKFZ High-Throughput Sequencing Core Facility to collect a median of 75 million reads per sample (spread: 65–99 million). Processing of RNA-Seq data was performed using a customized RNAseq Workflow as a part of the DKFZ—ODCF Roddy pipeline (https://github.com/DKFZ-ODCF/RNAseqWorkflow; version 1.3.0). After initial QC, the RNAseq raw reads were aligned to a customized hg19 version of the 1000 genomes reference genome (1KGRef_PhiX): (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/phase2_reference_assembly_sequence/hs37d5.fa.gz), by a 2 pass alignment using STAR aligner [34] (version 2.5.3a) and gencode 19 gene models. PhiX-genome sequence was added to the reference as an additional contig to remove spike-in sequences that had been included to find potential cross-contaminations and sample swaps. Additional QC after alignment was performed using samtools *flagstat* command and the RNA-SeQC tool [35] (version 1.1.8). The featureCounts function [36] as part of the SUBREAD software package [37] (version 1.5.1) was used to determine the respective raw counts per gene, using gencode 19 gene models. Downstream RNASeq analysis was performed with R (R version 4.1.0)/BIOCONDUCTOR software. The R package EnsDb.Hsapiens.v75 was utilized to annotate gene symbols. Non-specific intensity filtering was performed to exclude genes with less than 10 read counts in less than five samples. TMM normalization was conducted using the EDGER R package [38] (version 3.36.0), followed by incorporating sample weights and increasing statistical power using the Voom function. Finally, the R package immunedeconv [39] (version 2.1.0) was utilized to estimate the composition and density of immune and stromal infiltration, using the non-log-transformed Transcripts Per Kilobase Million (TPM) values as input.

2.7. TCGA and CPTAC multi-OMICs data

Public data from two CRC cohorts were downloaded from the cBioPortal multi-omics cancer portal (https://www.cbioportal.org/) that is, the PanCancer Atlas TCGA Colorectal adenocarcinoma [11] (https://www.cbioportal.org/study/summary?id=coadread_tcga_pan_can_atlas_2018) and the CPTAC Prospective Colon (https://www.cbioportal.org/study/summary?id=coad_cptac_2019) [40,41] datasets. Only samples were processed that had been profiled for SNVs, CNAs, and gene expression (RNAseq). Phosphoproteomic data were available only from CPTAC. SNVs were retrieved, and data analysis was performed using the R package maftools [42] (version 2.10.0). The significance of differential mutational frequencies between the MSS and MSI tumors within TCGA and CPTAC cohorts was assessed using Fisher's exact test based on a 2×2 contingency table (adjusted *P*-value ≤ 0.1, minimum number of samples harboring mutated genes ≥ 4), exploiting the mafCompare function of maftools.

2.8. Inference of microsatellite instability status in the CRC datasets

We utilized the MSI MANTIS [43] estimators of microsatellite instability provided with the TCGA pancancer dataset, using the following categorization: A TCGA tumor was classified as MSI if the respective value from MSI MANTIS was higher than 0.6, while tumors with relative values lower than 0.4 were classified as MSS. Any tumors with relative values between 0.4 and 0.6 were denoted as 'Undetermined' and excluded from further analysis. Information about MSI-status of respective samples was available also for the CPTAC dataset [40] and utilized as provided. Then, the PREMSIm R package [44] (version 1.0) was employed to infer the microsatellite status-based RNA-sequencing data from all three datasets (TCGA, CPTAC, ACCC), using non-log-transformed Transcripts Per Kilobase Million values from RNA-sequencing. MSI-L and MSS tumors were grouped together in all datasets and the term MSS used for this group to distinguish from MSI tumors.

2.9. Functional enrichment analysis

Functional enrichment analysis was performed using the R packages CLUSTERPROFILER [45] (version 4.2.2), and REACTOMEPA [46] (version 1.38.0) was used for overrepresentation analysis based on REACTOME molecular pathways [47]. SSGSEA [48] (https://github.

com/broadinstitute/ssGSEA2.0; accessed 2023/12/21) was implemented based on the SSGSEA projection methodology [49]. The hallmark signatures file (version 7.5.1) was used in gmt format from the respective Molecular Signatures Database (MSigDB) [50] gene sets (https://www.gsea-msigdb.org/gsea/msigdb/). The script ssgsea-cli.R (https://github.com/broadinstitute/ssGSEA2.0; accessed 2023/12/21) was run in a Windows 10 x64 (build 19 044) environment using command line and the following arguments:

$[-i \ ACCC.CRC.RNASeq.VoomCPM.TMM.Filtered$ $-o \ ACCC \ -n \ none \ -d. \ |ssGSEA2.0-master|db|$ $msigdb|h.all.v7.5.1.symbols.gmt \ -w \ 0.75 \ -c \ z.score \ -p$ $1000 \ -m \ 10]$. Protein associations were retrieved and visualized using STRING [51] (https://string-db.org; accessed 2023/12/04), and K-means clustering was performed to identify functional protein networks.

2.10. Pathway, TF, and kinase activity inference analyses

The upper-quartile of processed and log2 transformed RSEM estimated counts from CPTAC RNAseq data were normalized using the function *normalizeQuantiles* from the LIMMA R package [52] (version 3.50.0). Then, a non-specific intensity filtering procedure was implemented to remove features that were not expressed at least in the group with the lowest number of samples. The same preprocessing approach was applied to the TCGA dataset, with two additional steps: as the estimated raw RSEM counts had been batch effect-corrected, many features assigned by NA values were removed. Then, the data were log2-transformed if the maximum value was > 50. Differential expression analysis (MSI vs. MSS) was applied with the LIMMA R package [52] for both datasets.

The PROGENy (Pathway RespOnsive GENes R package) [53] (version 1.16.0) was implemented to identify putative differentially activated molecular pathways, based on the derived differential moderated statistics extracted from the LIMMA pipeline (i.e., top 100 most responsive genes for each pathway). PROGENy infers the activity of 14 major signaling pathways, based on the exploitation of consensus gene signatures and estimated from a large compendium of perturbation experiments [53]. The DOROTHEA-DECOUPLER pipeline [54] (version 1.6.0) was implemented to elucidate differentially activated regulatory networks. Relative activities were inferred for every transcription factor (TF) from the deregulated expression of its respective target genes, using the normalized weighted mean statistical metric from the DECOUPLER R package [55] (version 2.1.6). Customized barplots were

created using the R package GGPLOT2 [56] (version 3.3.5) for the visualization of differentially altered TFs and pathways.

Phosphoproteomics data from the CPTAC cohort were processed and analyzed to infer differentially activated signaling cascades between MSS and MSI colon tumors. To this end, normalized phosphosite ratios were used, representing scaled intensity values. Duplicated entries were removed and any phosphosites were excluded that had NA values across more than 85% of all samples. Differential protein phosphorylation analysis was performed using the LIMMA R pipeline (version 3.50.0). OMNIPATH-DECOUPLER [55] and Post-translational Modification Set Enrichment Analysis (PTM-SEA) [57] with the PTMSigDB signaling pathways' collection (version 2.0.0) were implemented to determine differentially activated kinases and signaling pathways, respectively. The OMNIPATH R package [58] (version 3.2.8) was utilized to retrieve prior knowledge interactions composed by kinase-target relationships, whereas the R package DECOUPLER [55] (version 2.1.6) was used to estimate the relative biological activities by application of statistical methods (normalized weighted mean). Along these lines, four filtering steps were applied to fetch and filter kinase-substrate interactions from Omnipath: (a) Signed enzyme-substrate interactions were fetched by querying the OMNIPATHR R package. (b) As OMNIPATH might contain putatively 'wrong/biased' records, such were removed that were only present in ProtMapper and not confirmed by other resources. If an interaction was from ProtMapper but also confirmed by another resource, it was retained. This step was performed because ProtMapper uses literature mining algorithms (REACH, Sparser), which are likely to produce false positives in their output [59]. (c) Only post-translational modifications were kept that involved dephosphorylation or phosphorylation events. (d) Duplicate interactions were removed based on the unique feature-interaction IDs.

2.11. Patient clustering

Unsupervised hierarchical clustering on principal components (HCPC) was implemented, based on the R packages FACTOMINER [60] (version 2.8) and FACTOSHINY (http://factominer.free.fr/graphs/factoshiny.html; version 2.4). The selected scaled pathway and TF activities per patient were used as input for the *PCA* function. Then, hierarchical clustering on the first 8 components (accounting ~ 95% of retained information) was performed using the *HCPC* (kk = Inf, max = 10, nb.clust = −1, consol = TRUE) function. Plots were created using the COMPLEXHEATMAP R

package [61]. Patients were re-ordered within each cluster according to their microsatellite instability status and selected mutational groups, namely BRAF_-mut, KRAS_mut, NRAS_mut, GNAS_mut, RAS_RAF_GNAS_mut, and RAS_RAF_GNAS_wt.

2.12. Identification of actionable alterations for construction of personalized cancer patient maps

Aiming to identify actionable vulnerabilities at the single patient resolution within the ACCC-CRC cohort, the qualified somatic variants from the SVRACAS scoring scheme were integrated with clinical implications from the OncoKB precision oncology database [62]. Variant oncogenic effect annotation was applied to keep only variants characterized as oncogenic or likely oncogenic. Therapeutically actionable variants in any cancer type were assigned with an OncoKB Therapeutic Level of Evidence (https://www.oncokb.org/therapeutic-levels; accessed 2022/12/21). Genes carrying actionable variants were assigned to the treatment baskets of the National Center for Tumor Diseases/German Cancer Consortium (NCT/DKTK) MASTER (Molecularly Aided Stratification for Tumor Eradication Research) program, that are based on a set of 472 genes which participate in biological processes and cellular pathways with established therapeutic recommendations [63].

3. Results

3.1. The mutational landscapes of MSS and MSI CRCs

To elucidate the landscape of somatic mutations in CRC, whole exome sequencing (WES) as well as RNA sequencing (RNAseq) was performed in cancerous as well as in adjacent normal tissues from 28 patients of a prospective patient cohort of the Athens Comprehensive Cancer Center (ACCC) (Table S1). Since microsatellite instability is one of the major clinical factors determining therapeutic management and prognosis of CRC patients, we first wanted to infer the microsatellite stability status of every patient using RNAseq data. To this end, we used PreMSIm [44] to predict MSI status in 434 patients of the TCGA Pan-Can COAD/READ [11] as well as in 79 patients of the CPTAC [64] cohorts. PreMSIm estimates were highly concordant with MSI-MANTIS predictions from TCGA (concordance rate: 0.988, Table S2), and with the annotated MSI-status of patients in CPTAC (concordance rate: 0.987, Table S3). We then

characterized the ACCC CRC samples applying PreMSIm and inferred five tumors as MSI and 23 as MSS (Table S4).

The exomes of ACCC CRC samples were sequenced at an average coverage of > 300× (range: 195–375). In total, 25 913 unique protein-altering somatic variants were identified in 11 492 genes, including 21 917 missense, 1495 stop_gained, 423 splice site, 1871 frameshift and 162 inframe InDel mutations (Table S5). The number of somatic variants varied greatly among ACCC patients as the tumor mutational burden (TMB) ranged from 0.028 to 375.994 protein-altering mutations per Megabase (Table S4). The median was 108 nonsynonymous SNVs in MSS tumors while that in MSI tumors was 1257. Five patients carried hot-spot mutations in the G12 codon of KRAS [G12D ($n = 3$), G12V ($n = 1$), G12S ($n = 1$)], another five patients carried hot-spot mutations in BRAF [V600E ($n = 4$), K601E ($n = 1$)], while concomitant mutations in KRAS [A146V] and BRAF [D594N] were observed in one ACCC-patient. Another patient carried an HRAS [A146T] mutation co-occurring with two non-hotspot mutations in BRAF [L505F and L577I]. The HRAS [A146T] substitution has been found as a germline variant in Costello's syndrome [65] and as a somatic mutation in a few cases of melanoma and other skin cancers (https://www.cbioportal.org, GENIE Cohort v14.1-public). BRAF [L505F] has been shown to confer a RAS-independent constitutively activated state in the kinase domain [66]. This and BRAF [L505H] have been reported in a small number of tumor entities including melanoma, prostate, and pancreatic cancer in cbioportal. BRAF [L505H] was shown to be activating mutation conferring resistance to vemurafenib [67], and is characterized as oncogenic in OncoKB [62]. The BRAF [L577I] mutation has been reported in cutaneous melanoma and is characterized as variant with 'Unknown Oncogenic Effect' in OncoKB.

In order to prioritize somatic mutations according to their putative oncogenicity, we next applied our SVRACAS variant scoring pipeline (https://github.com/Jasonmbg/Simple.-Variant-Ranking-Annotation-CAncer-Score) to the 25 913 variants, leaving 4307 prioritized variants (score ≥ 0.5) in 2969 genes (Table S5). Among these, 54 matched with 72 CRC driver genes annotated in IntOgen [30] (Fig. 1). The top frequently mutated genes in the ACCC tumors were *APC* and *TP53* (Fig. 1). Next, we revealed overlaps and differences in the 54 mutated genes between the MSI and MSS CRC tumors of the ACCC (Fig. 1) cohort regarding somatic SNVs, InDels and copy number alterations (CNAs). On the one hand, MSS
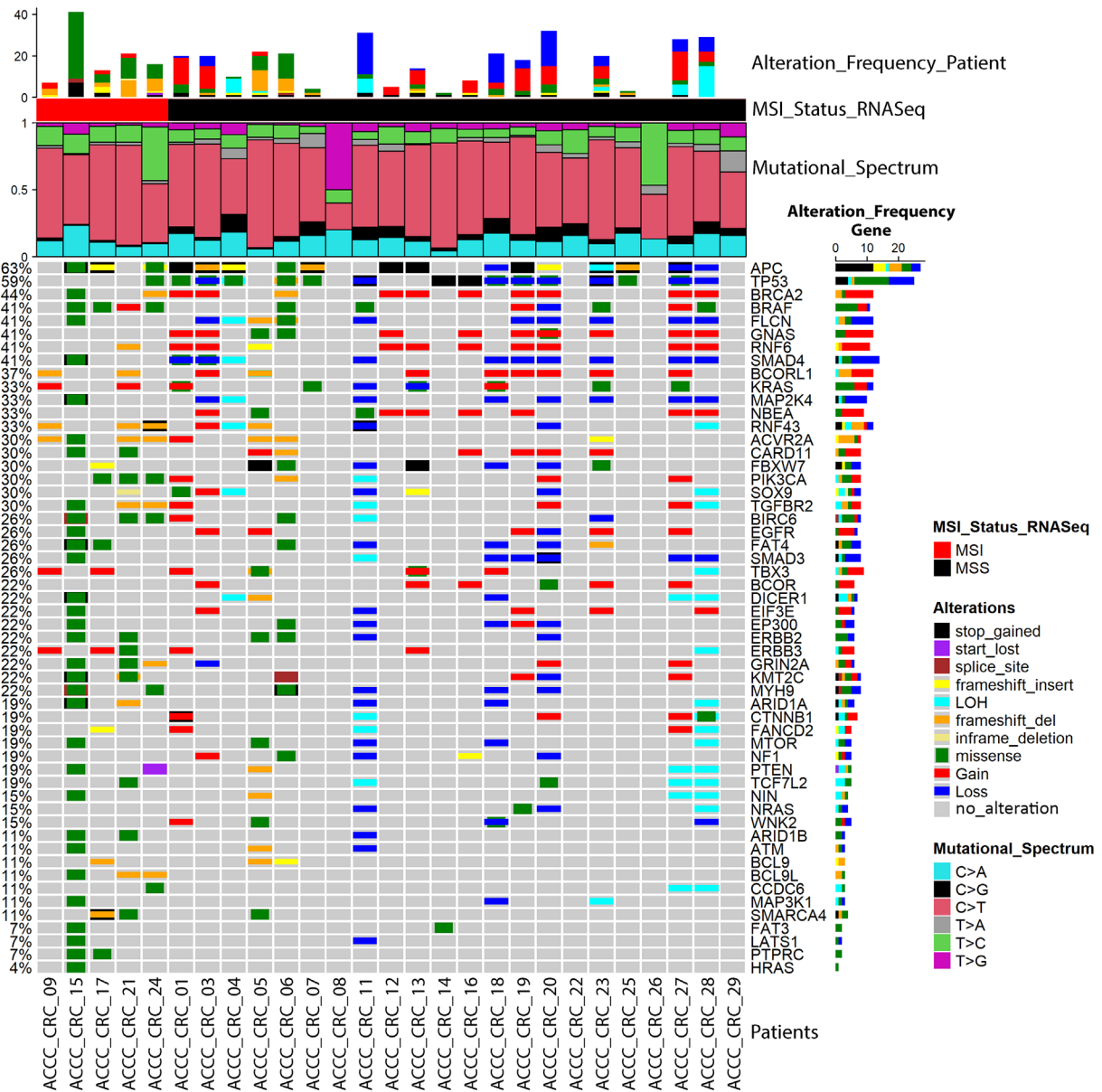
Fig. 1. Mutational landscape in the tumors of the Athens Comprehensive Cancer Center Colorectal Cancer (ACCC-CRC) cohort. Oncoplot presenting scored genes common with the IntoGen database for 27 ACCC-CRC patients, which were grouped into microsatellite instable (MSI) and microsatellite stable (MSS) subtypes (i.e., MSI_Status). Each column represents a profiled patient and each row a gene of interest. The alteration frequencies of 54 genes (single nucleotide variants (SNVs), small insertions/deletions (InDels) and copy number variations (CNVs)) are shown for each patient and for each gene on the top and on the right barplots, respectively. Genes are sorted based on their total mutation frequency (depicted on the left). Colored cells indicate respective types of alterations explained on the right, while gray cells indicate lack of mutations in respective tumors. The mutational spectrum for each patient showing the relative contribution of each of the six base substitution types is presented below the top bar plot as a stacked bar chart, using the Mutational Patterns R package [116] (version 3.4.1).

tumors consistently exhibited a higher percentage of CNAs: Deletions of *SMAD3*, *SMAD4* and *MAP2K4*, and gains in *RNF6*, *BRCA2*, *NBEA* as well as in *GNAS* were only observed in MSS tumors. On the other hand, frameshift and missense variants were more frequent (*RNF43*, *ACVR2A*, *TGFBR2*) or exclusive (*BCL9L*, *PTPRC*) in the MSI tumors. These findings are in line with reports linking defective DNA

mismatch repair, frameshift mutations, the MSI sub-type, and the success of immunotherapies [68].

Mutational analysis implementing the COSMIC SBS signatures [23] revealed distinct patterns between the MSI and MSS tumors (Fig. S1). The majority of the MSI tumors were characterized by SBS6, SBS26, and SBS20 signatures, all related to defective mismatch repair and, thus, microsatellite instability. One sample (ACCC_CRC_15) was further characterized by signatures 10a and 10b, which are related with mutations in *POLE*. Mutations in *POLE*, as was also present in this tumor, are associated with a very high TMB, which was indeed observed in this patient (375.994 mutations/Mb). On the contrary, MSS samples were mainly enriched in the SBS1 and SBS5 signatures, which are related to age, stem-cell division and 'clock-like' processes, and are in line with a generally low TMB in MSS tumors. However, the DNA mismatch repair-associated signatures SBS6 and SBS15 were evident also in two MSS cases (ACCC_CRC_05 & ACCC_CRC_06, respectively), possibly related to their high TMB and mutations in several genes involved in DNA mismatch repair (e.g., *MLH3*, *MSH3*, *MSH6*). Patients ACCC_CRC_07 and ACCC_CRC_18 were enriched with the unknown etiology signatures SBS93 and SBS94, which have been identified in gastric and colorectal samples, respectively. Taken together, these findings are in line with previous studies recapitulating the mutational landscapes of CRC, where microsatellite instability is associated with a low frequency of CNAs and a high TMB [69,70].

3.2. Mapping mutational landscapes to functional networks

To extend the analysis of mutational profiles in MSS and MSI CRC tumors, we utilized the public datasets of the TCGA PanCan COAD/READ [11] and the CPTAC [64] cohorts. Complete datasets were available for 434 patients (392 MSS, 42 MSI) from TCGA (Table S2), and 79 patients (65 MSS, 14 MSI) from CPTAC (Table S3). Highly scoring variants from SVRACAS annotation (score ≥ 0.5) were identified in 7287 and 2588 genes in the TCGA and CPTAC datasets, respectively (Tables S6 and S7). Of those, only three genes had higher mutation frequencies in MSS compared to MSI tumors: *APC* and *KRAS* genes were found more frequently mutated in MSS in both public datasets (Figs S2 and S3), whereas mutations in *TP53* were significantly enriched only in the TCGA PanCancer dataset. In contrast, 564 (TCGA) and 80 (CPTAC) genes had significantly higher mutation frequencies in MSI compared to MSS tumors, and 42

genes were shared in all three datasets. To uncover molecular networks possibly related to microsatellite instability, we then performed functional enrichment analysis by mapping the genes differentially mutated (MSI vs. MSS) in TCGA and CPTAC to Reactome [47]. Eleven pathways were significantly enriched in MSI vs. MSS tumors in both TCGA and CPTAC. These signaling pathways were related to chromatin modification and organization, growth factor, second messenger as well as Wnt signaling, and to RUNX1 transcriptional regulation (Fig. 2). No Reactome pathway was enriched in MSS tumors, in line with only three genes having higher mutational frequencies in MSS vs. MSI.

Next, we mapped the 2969 mutated genes from the ACCC CRC dataset to the 11 enriched Reactome pathways. This identified 60 genes that were mutated with a higher frequency in ACCC MSI tumors and in at least one of the two other CRC datasets. Functional and physical associations were investigated by mapping these genes to the STRING database [51], identifying three protein interaction networks with 60 genes and 154 edges (Fig. S4). The first network was composed of genes in angiogenesis as well as ERBB2, PI3K/AKT, and PDGFR signaling. Network 2 was enriched in histone modifications, NOTCH1, TCF, and Wnt/β-Catenin signaling, while network 3 was characterized by RUNX1 signaling, regulation of nucleotide-excision repair, double-strand break repair, and chromatin remodeling. The Reactome and STRING network topologies not only mirrored prior knowledge of MSI CRC, thus further validating the ACCC CRC cohort and the classification of patients into MSS and MSI subtypes [11] but they also revealed significant functional networks that are differentially disrupted between MSS and MSI CRC tumors. These were driven by their distinct mutational profiles and, using this approach, could be uncovered independent of the size of the respective cohort.

3.3. Molecular pathway and transcription factor activities are distinct in MSS vs. MSI mutated CRCs

Aiming to estimate the activity of cancer-relevant signaling pathways in MSS and MSI groups of the three CRC datasets (TCGA PanCan COAD/READ [11], CPTAC [64] and ACCC CRC), we next estimated the activities of 14 signaling pathways based on transcriptomic data using the PROGENy computational pipeline [53]. The activity only of the TGFβ signaling pathway was significantly higher in MSS than MSI tumors in all three cohorts (Fig. 3A), which is in line
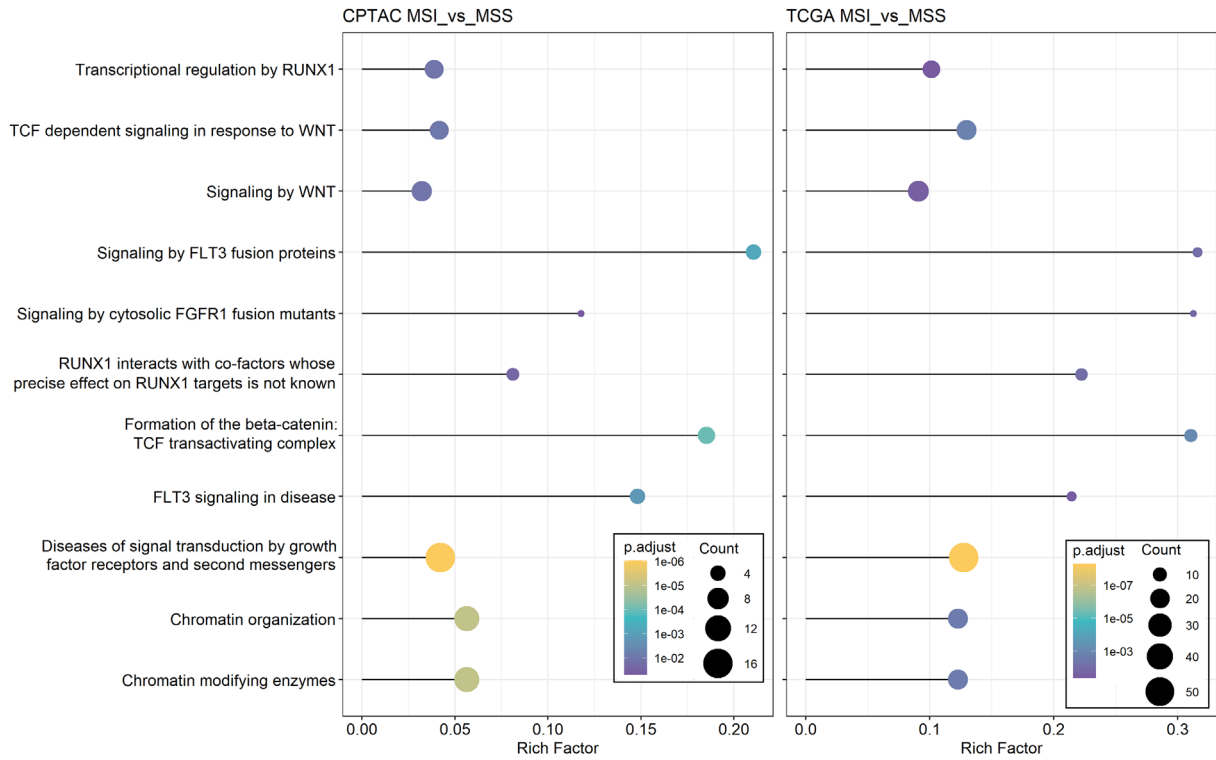
Fig. 2. Lollipop plots of REACTOME pathways significantly enriched in microsatellite instable vs. microsatellite stable tumors of The Cancer Genome Atlas Network (TCGA) and National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC) cohorts. Functional enrichment analysis was performed using CLUSTERPROFILER [45] (v.4.2.2), and GGPLOT2 (v.3.4.3) was used for the visualization of enrichment results. *P*-values were adjusted for multiple comparisons (*P*.adjust) [117]. Sizes of bullets represent the number of over-represented genes (Count) in the corresponding pathway and adjusted *P*-values (*P*.adjust) are color-coded. The Rich Factor value (*x*-axis) describes the relative ratio of the enriched genes in each pathway, divided by the total number of annotated genes in the respective set (i.e., over-representation analysis) [117].
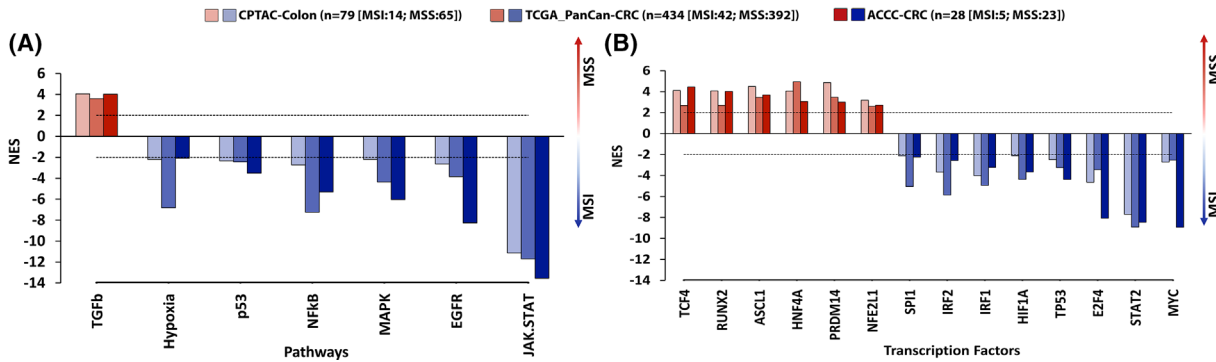
Fig. 3. Significantly differentially activated pathways and transcription factors (TFs) between microsatellite instable (MSI) and microsatellite stable (MSS) colorectal cancer (CRC) in three CRC cohorts; National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC), The Cancer Genome Atlas Network (TCGA), and Athens Comprehensive Cancer Center Colorectal Cancer (ACCC). Pathway and transcription factor (TF) relative activities were estimated using PROGENY [53] and DOROTHEA [54] R packages, respectively. Barplot visualizes normalized Enrichment Scores (NES) with three different shades of red and blue color denoting significant pathway/TF activity levels of MSI and MSS tumors in CPTAC, TCGA, and ACCC CRC datasets, respectively. Horizontal dotted lines at +2/−2 values indicate the lower cut-off values for statistically significant activities of pathways (A) and TFs (B). Only pathways and TFs found significantly differentially activated in all three datasets are illustrated. The total numbers of samples characterizing each CRC group (MSI or MSS) are denoted in parentheses.

with reports suggesting that TGFβ signaling in MSS tumors stimulates cancer-associated fibroblasts and establishment of an immune-excluded tumor microenvironment [71]. In contrast, six pathways were stronger activated in MSI tumors of all three cohorts (Fig. 3A). JAK–STAT signaling was the most activated pathway in MSI and has been implicated with positive immunotherapy outcomes [72]. Then, kinase and pathway activity analyses performed using the phosphoproteomics measurements from the CPTAC cohort. This suggested the cAMP-dependent protein kinase and TGFβ signaling pathways as being more active in MSS tumors compared to MSI (Fig. S5A,B), confirming the PROGENy results for TGFβ. Activation of protein kinase A might be functionally related to gains and correlated expression of *GNAS* ($r = 0.45$, *P*-value < 0.05) found specifically in MSS tumors also in the CPTAC cohort (Fig. S5C,D). This pathway has been proposed as a targetable driver mechanism in lung cancer [73]. In contrast, several MAP kinases and PI3K signaling were significantly more activated in the MSI tumors (Fig. S5A,B). The latter findings confirm that MSI tumors exhibit a higher activation of stress response, inflammation/immune modulation, and proliferation. Yet, differential activation of these pathways was not common in all MSI patients, suggesting alternative driver routes in respective tumors (Fig. 4, and Figs S6 and S7).

Next, we performed transcription factor (TF) activity analysis in MSS vs. MSI tumors and found 14 TFs that had consistently different activity profiles in all three datasets. Six transcription factors including RUNX2 (Runt-related transcription factor) were more strongly activated in MSS CRCs compared to MSI in all datasets (Fig. 3B). RUNX transcription factors are downstream effectors of TGFβ, thus matching the results from pathway analysis. These TFs regulate basic cellular and developmental processes, stem cell biology, and tumorigenesis [74]. In colorectal cancer, RUNX2 influences MAP kinase signaling via regulation of multiple RTKs, and its absence, can lead to resistance against MEK inhibitors along with its cofactor CBFB [75]. Furthermore, MSI CRCs presented significantly higher activities for the MYC, STAT2, HIF1A, TP53, SPI1, IRF1, and IRF2 transcription factors (Fig. 3B). Some of these have previously been associated with CRC and potentially impact immunogenicity, proliferation, DNA repair mechanisms, and cellular stress responses [76]. Collectively, pathway and transcription factor activation-profiling point at elevated inflammatory processes in many MSI tumors while tumors of the MSS type mostly appear to be associated with stem-cell features as well as an immune-excluded microenvironment.

3.4. Transcriptomic stratification and immune deconvolution of ACCC-CRC samples

We then estimated altered pathway and TF activities at the individual patient level to achieve a personalized view on the ACCC patients. To this end, we used the scaled gene expression values of every patient as input and focused on the seven molecular pathways and 14 TFs, which we had found significantly altered in the three cohort-based comparisons of MSS vs. MSI CRC tumors. The resulting estimated pathway and TF activities grouped the CRC patients into three patient clusters in each cohort (Fig. 4A, and Figs S6 and S7). While we had observed elevated TGFβ-signaling specifically in the MSS subtype, analysis of individual patients revealed that this pathway clustered specifically with MSS tumors that had reduced activation of JAK/STAT, MAPK, and EGFR signaling (Fig. S6). Furthermore, few MSI tumors showed strong activation of TGFβ signaling as well (Figs S3 and S4). To gain further insights into the underlying mechanisms that characterize the identified clusters, we performed single sample Gene Set Enrichment Analysis (ssGSEA) [48], immune deconvolution [39], and inference of Consensus Molecular Subtypes (CMS) [77].

The first patient cluster (Cluster_ID 1, $n = 13$ patients in the ACCC cohort) mostly included MSS ($n = 12$) and CMS2 and CMS3 cancers (Fig. 4A). In line with MSS tumors in the TCGA and CPTAC (Fig. 3A), the tumors in cluster 1 were characterized by a low immunogenic behavior with sparse activation of the JAK–STAT and NFkB inflammatory pathways, no enrichment of immune-related hallmarks, and flat immune infiltration scores (Fig. 4B). These findings are in line with the poor response MSS patients mostly have to immunotherapies. Patient ACCC_CRC_15, the only MSI sample in this cluster, also showed an overall trend of under-activation in the majority of estimated pathway and TF activities, whereas the EGFR signaling cascade and MYC TF were found significantly over-activated in this *POLE*-mutant tumor. This tumor had been diagnosed as serrated mucinous adenocarcinoma, which might explain the separate clustering from the other MSI tumors [78]. The second cluster (Cluster_ID 2, $n = 7$) contained only MSS and mostly CMS4 samples ($n = 6$) which were characterized by an overall significant activation of the TGFβ signaling pathway, expression of genes associated with epithelial to mesenchymal transition (EMT) and angiogenesis as well as activation for the SPI1, RUNX2, and PRDM14. Almost all samples in the second cluster exhibited low activity of E2F4 and
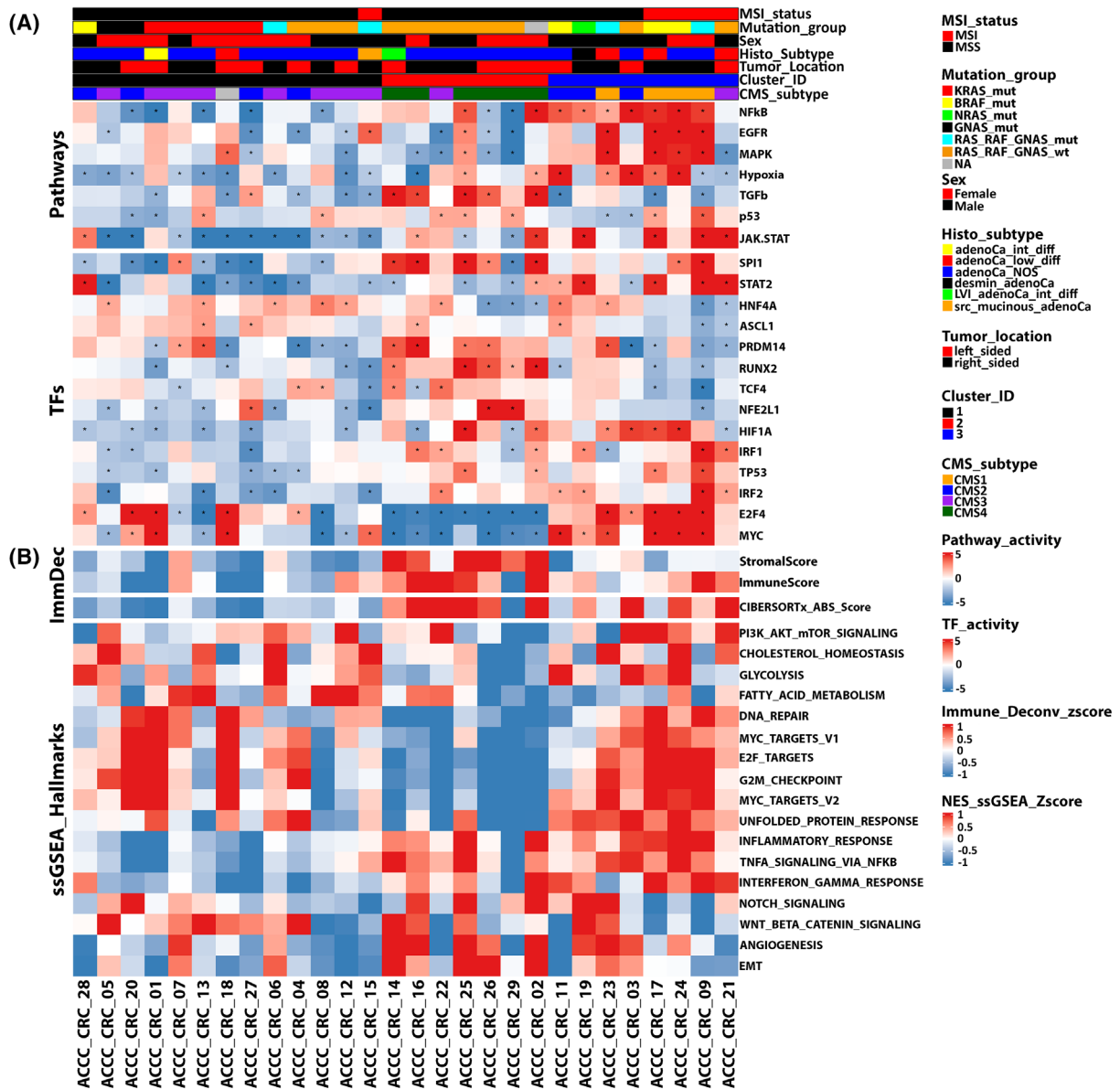
Fig. 4. Clustering of Athens Comprehensive Cancer Center Colorectal Cancer (ACCC-CRC) patients based on transcriptomic analysis. (A) Normalized enrichment scores (NES) of 7 pathway and 14 transcription factor (TF) activities (compare Fig. 3) were used for unsupervised hierarchical clustering on principal components (HCPC) with R packages FactoMineR [60] (version 2.8) and FACTOSHINY (http://factominer.free.fr/graphs/factoshiny.html; version 2.4), distributing ACCC patients into three patient clusters. Features were re-ordered within each Pathway and TF row sub-clusters using hierarchical clustering ('complete' method, 'euclidean' distance), where samples were firstly grouped based on their cluster ID and subsequently based on their MSI status and mutational information. Tumors with co-occurring alterations in either in *RAS* and *RAF* or *GNAS* and *RAF* genes were assigned to the RAS_RAF_GNAS_mut group. Pathway and TF activities (NES) are indicated in shades of red (high) vs. blue (low). Asterisks in individual cells denote a significant activation of the respective pathway/TF in a patient (NES $\geq$ 2 or NES $\leq$ −2). (B) Heatmap visualization of the normalized enrichment scores of hallmarks from MSIGDB [50] per patient using SSGSEA [48]. Re-ordering of rows was performed as described in (A). Hallmark pathway activities as well as immune or stromal scores (ImmDec: immune deconvolution), representing immune and stromal cell infiltration in respective tumors, are indicated in shades of red (high) vs. blue (low). Plots were created using the COMPLEXHEATMAP R package [61].

MYC, accompanied with reduced cell cycle activity and unfolded protein response (Fig. 4A,B). This suggests that tumors in the second cluster are characterized by an overall anti-proliferative capacity and a more invasive phenotype, which could be related to the molecular and clinical characteristics of the

mesenchymal subtype (CMS4) [79]. The third cluster (Cluster_ID 3, $n = 8$) comprised four MSI and four MSS samples. It was mainly composed of CMS1 and CMS2 tumors with the exception of sample ACCC_CRC_21 that was predicted as CMS3. Tumors in cluster 3 were characterized by a relatively stronger activation of inflammatory (JAK–STAT, NFkB) and mitogenic (EGFR and MAPK) signaling pathways, compared to patients in the other clusters. Moreover, MYC and E2F4 seemed to be strongly activated. Conversely, a group of TFs, mainly associated with chromatin remodeling (PRDM14, ASCL1, RUNX2) and developmental processes (HNF4A, TCF4) displayed a reduced activity level compared to the other two clusters. This was most evident in the MSI samples of the third cluster, while MSS samples in this cluster showed a more variable activation pattern. Collectively, patients forming the third cluster reflected a more immune-reactive and less aggressive molecular profile.

The MSS tumors showed an overall higher degree of heterogeneity, which was associated with the prevalence of *KRAS* and *BRAF* mutations. *KRAS*-mutated MSS samples showed a trend toward having lower activation of immune response pathways, that is, having significantly lower activities in JAK–STAT and NFkB pathways relative to the RAS_RAF_GNAS_wt MSS samples in the second patient cluster (Fig. 4A). This was corroborated by ssGSEA findings and immune deconvolution estimates (Fig. 4B), suggesting that *KRAS*-mutant MSS tumors in the first patient cluster were less immune reactive compared to the MSS tumors in the other clusters. These findings are in line with the observed low activity of IRF1 and IRF2, regulators of the interferon response [80], in the majority of the *KRAS*-mutant MSS samples of the first cluster (Fig. 4A). This is consistent with previous studies, which identified a mutant *KRAS-IRF2* molecular cascade, promoting an immunosuppressive environment in CRC and fostering immunotherapy resistance [81]. Conversely, the transcriptomic circuits of MSS CRC tumors in the third patient cluster showed greater similarity to the respective circuits in the MSI tumors within that cluster, compared to the MSS samples in the other clusters. In summary, while the evaluated transcriptomic programs primarily differentiated MSS and MSI samples and corresponded with the prevalence of identified CMS subtypes, they also exposed differentially enriched patient-cluster and individual patient-specific gene expression networks. Analysis of the tumors from the TCGA and CPTAC cohorts revealed highly congruent findings (Figs S3 and S4).

3.5. From oncogenic mutational landscapes and transcriptomic features toward personalized therapeutic strategies

Having seen higher functional heterogeneity in MSS compared to MSI tumors and that only a few pathways seemed to be consistently stronger activated in MSS tumors, we hypothesized that precision oncology approaches might be particularly useful for MSS CRC. Along these lines, we went out to unravel actionable alterations for individual ACCC CRC patients and interrelate them with their respective transcriptomic profiles. This should provide more holistic views on potential vulnerabilities and frame them as comprehensive personalized cancer maps, placing our special emphasis on MSS cases. For this purpose, the SVRACAS prioritized somatic variants in each patient were annotated with clinical actionability information from the OncoKB precision oncology database [62]. This analysis prioritized 34 actionable variants in 20 genes, reflecting these variants as predictive biomarkers of drug sensitivity or resistance to particular targeted therapies (Fig. 5A).

Next, we assigned the ACCC CRC patients to seven MASTER interventional baskets [63], based on the mutational profiles the respective tumors presented with. The MSI status and TMB were regarded for assignments to the immune evasion (IE) basket. Seventeen patients of the ACCC cohort (~ 63%) were stratified into at least one interventional basket (Fig. 5B). Tumors in patient cluster 1 were mainly characterized by overrepresentation of the RAS–MEK–ERK (RME, eight out of nine patients), the PI3K-AKT–mTOR (PAM, three out of nine patients), and the immune evasion (IE, three out of nine patients) baskets. An interventional basket (i.e., RME) could be assigned to just one (ACCC_CRC_16) out of seven MSS patients forming cluster 2, in line with the rather low TMB in the tumors forming this cluster. Cluster 3 was mostly represented by RME (five out of seven patients) and PAM as well as IE (four out of seven patients each) baskets.

The five MSI- and two MSS-patients could qualify for treatment with immune checkpoint inhibitors (e.g., pembrolizumab), the MSS-patients in the light of the high TMB in the tumors (Fig. 5B). Two MSI and another two MSS tumors carried BRAF V600E mutations and might benefit from treatment with encorafenib and cetuximab [82]. Seven patients carried *KRAS* or *NRAS* mutations, rendering those patients incompatible with EGFR-antibody therapy [83]. Inhibitors specific for the most common KRAS variant G12D,
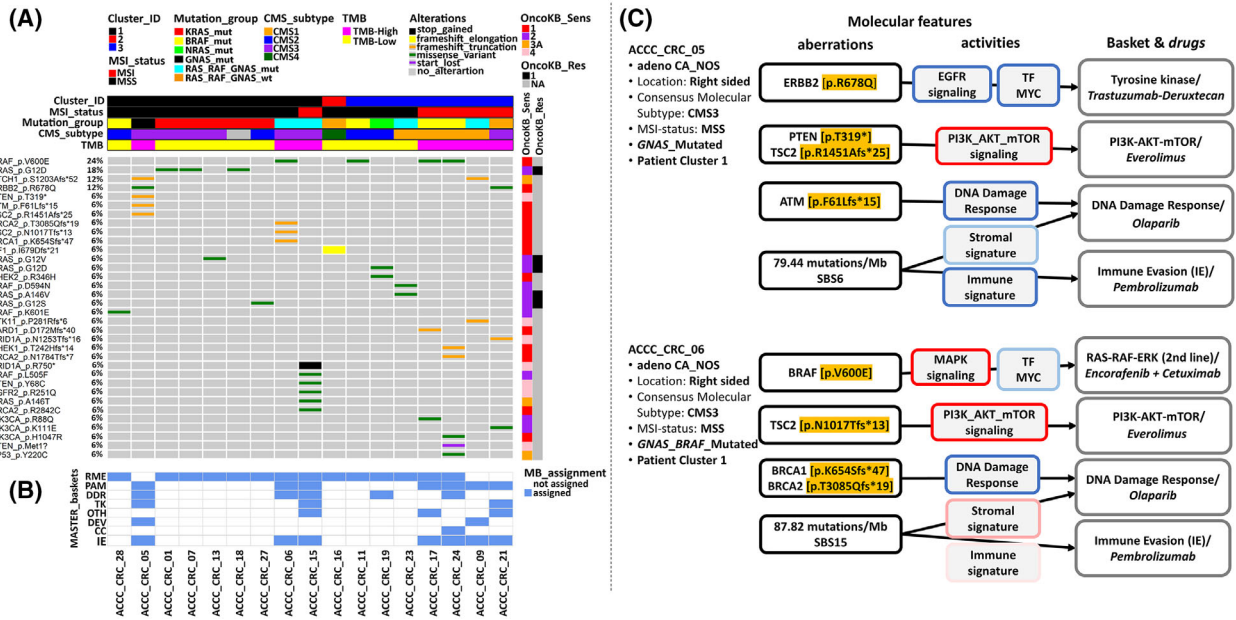
Fig. 5. SVRACAS prioritized somatic mutations, pathway/transcription factor (TF) activities, and immune profiles stratify Athens Comprehensive Cancer Center (ACCC) Colorectal Cancer (CRC) patients to MASTER baskets and personalized therapeutic interventions. (A) Oncoplot with actionable variants (SVRACAS score ≥ 0.5) and annotated in the context of sensitivity (OncoKB_Sens 1–4) or resistance (OncoKB_Res R1-2) to a targeted therapy (https://www.oncokb.org/levels) in any tumor entity. Columns correspond to patients (B) and rows denote genes and protein-level types of alterations. Types of alterations are indicated in color, while wild type condition is indicated in gray. Rows were sorted by the percentage of occurring alterations, whereas columns/samples were firstly grouped based on patient cluster ID, then by microsatellite instability status, and subsequently by selected driver genomic alterations (i.e., KRAS/BRAF/RAS_RAF_WT). TMB: Tumor mutational burden. Tumors with co-occurring alterations in either in RAS and RAF or GNAS and RAF genes were assigned to the RAS_RAF_GNAS_mut group. (B) Mapping of indicated patients to MASTER interventional baskets (MB_assignment in blue) based on mutated genes that are associated with respective baskets [115] (RME: RAF–MEK–ERK; PAM: PI3K–AKT–mTOR; DDR: DNA damage repair; TK: tyrosine kinases; OTH: other; DEV: developmental regulation; CC: cell cycle). Mapping to the immune evasion (IE) basket was based on the presence of MSI and/or high TMB. MB_Assignment in white: not assigned. (C) Molecular rationale for suggested personalized therapeutic interventions for two ACCC_CRC patients. The first series of boxes (aberrations) illustrates actionable variants (in orange) and TMB as well as microsatellite instability (MSI) related COSMIC single base substitution (SBS) signatures [23]. The second series (activities) presents perturbed molecular pathways, TF-MYC, and stromal and immune signatures, with boxes framed in shades of red or blue to indicate either high or low activity levels and high or low scores, respectively. The final series (Baskets & drugs) shows interventional baskets [118], with suggested drugs, based on a comprehensive analysis of actionable genomic variants, molecular pathway/TF activities, and immune profiles.

which was present in three cases, are currently in pre-clinical and clinical development (e.g., MRTX1133, NCT05737706). Tumors carrying druggable KRAS G12C mutations [3] were not present in our cohort. No specific treatments were clinically available for the patients carrying other RAS mutations, that is, ACCC_CRC_13 (G12V) and ACCC_CRC_27 (G12S). However, also non-specific Pan-KRAS inhibitors are in development [84], which might be applied once having proven efficacious. Hence, molecular profiling identified potential treatment options for the majority of the patients in the ACCC cohort and more than one interventional basket was mapped to eight patients, suggesting a combination of targetable driver mechanisms in the respective patients.

We finally went ahead with two patients of the MSS type to prove the utility of our approach toward suggesting potential vulnerabilities for tailored therapeutic interventions. Patients ACCC_CRC_05 and ACCC_-CRC_06 had presented with adenocarcinomas of the CMS3 subtype and had a right-sided location. Patient ACCC_CRC_05 (Fig. 5C) was characterized by an ERBB2 p.R678Q mutation, which has a modest biological rationale for treatment with trastuzumab-deruxtecan [85] (NCT evidence level m4). Furthermore, this patient carried an ATM p.F61Lfs*15 mutation, which has been found in few, mostly colorectal tumors (CosmicID: COSV53735143). While preclinical evidence supports PARP inhibition [86] (NCT evidence level m3) this was recently shown (NCT02693535) to

be ineffective in *ATM*-mutated CRC [87] thus not giving it high priority. In patient ACCC_CRC_05, we identified *PTEN* p.T319* and TSC2 p.R1451Afs.25 mutations, both of which have biological justification and preclinical evidence [88] suggesting treatment with mTOR inhibitors (NCT evidence level m3-m4). Tumor gene expression analysis for this patient demonstrated an over-activation of PI3K/AKT/mTOR signaling. The data further suggested an under-activation in the EGFR and JAK–STAT signaling pathways, which might indicate a predominantly PI3K/mTOR-driven cancer phenotype for this patient, further supporting the rationale for mTOR inhibition.

MSS patient ACCC_CRC_06 carried several potentially druggable alterations (Fig. 5C). The BRAF V600E mutation in this tumor could be targeted with Encorafenib and Cetuximab as an approved combination for second-line targeted therapy after chemotherapy [2]. BRCA1 p.K654Sfs*47 and BRCA2 p.T3085Qfs*19 mutations in this patient suggested PARP inhibition as a potential option. Tumor transcriptomic analysis revealed an over-activation of the PI3K/AKT/mTOR and under-activation of the DNA repair and JAK–STAT molecular circuits. A frameshift mutation in *TSC2* (p.N1017Tfs*13) would suggest mTOR inhibition as a therapeutic strategy [89]. However, in contrast to patient ACCC_CRC_05, we noticed a trend toward activation of the MAPK signaling pathway in patient ACCC_CRC_06, which might be associated with the presence of the BRAF V600E mutation. Thus, the value of targeting mTOR in the context of an activating BRAF mutation seems questionable [90].

Patients ACCC_CRC_05 & _06 both presented with high TMB (Fig. 5), which has been observed in MSS tumors before [91] and also in five patients (i.e., TCGA-AG-A002, TCGA-AZ-4315, TCGA-AA-3984, TCGA-EI-6917, TCGA-AA-3977) of the TCGA dataset (Table S2). The tumor DNA showed mutational signatures associated with DNA mismatch repair (SBS6 and SBS15 in patients ACCC_CRC_05 and ACCC_CRC_06, respectively). The two ACCC with high TMB carried mutations in a number of genes having roles in DDR and MMR pathways (*ATM*, *BRCA1*, *BRCA2*, *MSH3*, *MSH6*, *POLG*, *RAD50*), which were mutated also in some or all of the five TCGA MSS tumors. In addition, a number of epigenetic writers were mutated (*KDM4C*, *KDM5A*, *KMT2A*, *KMT2C*, *KMT2D*, *SETDB1*), again in several of the two ACCC and the five TCGA patients. The high TMB, which might be causally associated with defective damage response pathways, could indicate the application of immunotherapies as an at least

hypothetical treatment option for the two ACCC patients [92]. Yet, our analysis pointed also at other rational treatment options for these two as well as for most other patients in the ACCC CRC cohort. Collectively, our study underlines the significant value of incorporating various molecular and functional layers to support therapeutic decision-making.

4. Discussion

Here, we set out to characterize the molecular and activity landscapes of MSS and MSI colorectal cancer. By implementing an integrative computational workflow, we analyzed tumors from 28 patients of a prospective CRC patient cohort of the Athens Comprehensive Cancer Center (ACCC), in combination with two large public multi-omics datasets. There, we explored the molecular profiles of MSS vs. MSI CRCs, their relation to *KRAS*/*BRAF* mutations, and pathway as well as transcription factor activities in the context of personalized medicine. First, we observed novel associations between particular genes and their increased mutational frequencies specifically in MSI tumors. Increased tumor mutational burden (TMB) as well as elevated numbers of frameshift mutations are frequent in MSI and have been associated with defective DNA mismatch repair in this subtype as well as with response to immunotherapy [18,93]. Yet, we observed high TMB also in some MSS tumors, which has been reported before [91]. Mutations were enriched in genes that are involved in molecular pathways mainly related to transcriptional regulation cascades by RUNX1, chromatin modification mechanisms, and in signaling cascades that are related to Reactome 'fusion mutants/proteins' (i.e., FGFR1/FLT3). These findings resonate with research indicating the central role of RUNX1 in the malignant transformation process and metastasis of CRC through mechanisms, like the Wnt/β-catenin signaling pathway and EMT [94,95]. Fusion events involving FGFR have been considered as therapeutic targets in various solid tumors [96]. In contrast, MSS tumors mostly demonstrated a pronounced enrichment in CNAs. Deletions were recurrently found in *SMAD3*, *SMAD4*, and *MAP2K4*, which are pivotal in the TGFβ and JNK signaling pathways [97,98], respectively. Gains were observed in chromosomes 13q12-13 and 20q13.32, which have been associated with early events in the development of colorectal cancer [99], also affecting *GNAS*. These CNAs are exclusively presented in MSS, suggesting a unique MSS genomic signature with implications in MSS disease etiology and therapeutic targeting [100].

Second, our comprehensive multi-omics analysis indicated differentially regulated circuits at the transcriptional level. Immune-related biological mechanisms, such as JAK–STAT and NFKB signaling pathways, were found upregulated in MSI tumors, consistent with the higher 'immunogenicity' in the context of microsatellite instability [101]. Conversely, the TGFβ pathway showed greater activation in MSS tumors and has been associated with epithelial-to-mesenchymal transition (EMT), metastasis, and therapy resistance [102,103]. We further observed that the synchronous presence of *KRAS* mutations with microsatellite instability had a stronger effect on the immune profile of CRC tumors than mutations in *BRAF*. Hence, we identified context-specific functional associations that are likely related to the higher functional heterogeneity within MSS CRC tumors. Previous studies have emphasized the intricate role *KRAS* mutations play in shaping the tumor microenvironment of colorectal cancer. These mutations seem to foster an immunosuppressive environment, which is marked by a diminished anti-tumor immune response [104].

Next, we identified transcription factors with differential activity between MSS and MSI CRC tumors. MSI CRC tumors showed higher activities of MYC, STAT2, HIF1A, TP53, SPI1, IRF1, and IRF2. These factors are instrumental in various cellular processes, such as cell growth, response to stress, oxygen homeostasis, and immune modulation [76]. Hence, therapeutic regulation of TFs might be a promising strategy for CRC therapy owing to their underlying role as key effectors of various signal transduction and tumor-associated immune responses [105]. Distinct transcriptional profiles were observed in MSS CRCs, indicating the heightened activity of RUNX2 upstream of TGFβ signaling [106]. Yan et al. [107] revealed that RUNX2 is critical for the maintenance of the stem cell-like properties of CRC cells as well as for the promotion of CD44-induced EMT in CRC. Collectively, from our integrative analysis, we noticed the significant deregulation of specific members of the RUNX TF family at DNA (*RUNX1*) or mRNA (*RUNX2*) levels, supporting their emerging role as prognostic biomarkers in CRC [108].

Our comprehensive molecular tumor analysis thus corroborated prior findings related to the phenotypic characteristics of CMS subtypes; however, also suggests that the microsatellite status is not the only biomarker that could predict molecular and functional tumor characteristics. In our clustering analysis, one cluster grouped MSI with MSS patients together. This could be explained, in part, with epigenetic reprogramming in tumor progression, playing a major role in the

clonal evolution and tumorigenesis of CRC [109]. Indeed, this aligns with a recent study proposing a novel molecular classification of CRC tumors based on single-cell sequencing data [110]. This study identified a subgroup of MSS samples (iCMS3_MSS), which exhibited more similarity with the transcriptional and genomic levels of MSI CRC tumors (iCMS3_MSI), forming a novel epithelial subtype [110]. This may indicate that the MSI status, even though clearly distinct at the genomic level, may be viewed as a continuum through different states and sub-conditions at the transcriptomic layer [111]. The potential therapeutic consequences of this, however, remain to be uncovered.

Finally, we sought to provide a personalization of molecular findings, as the identification of targetable alterations is a primary focus of current clinical research, to facilitate personalized treatment strategies. Effective therapeutic options exist for MSI CRCs employing immune checkpoint inhibitors. In contrast, the management of MSS cancers still predominantly relies upon conventional chemotherapies, as the most frequently encountered mutations in MSS CRC affect yet non-druggable tumor suppressor genes, such as *APC* and *TP53*. Indeed, the prevalence of well-established targetable mutations with clinical evidence supporting therapeutic interventions is exceedingly low in CRC. These primarily encompass BRAF V600E mutations, which we observed also in our cohort and have proven more difficult to target in CRC compared, for example, to melanoma [82]. Clinical treatment options are often limited even when mutations in potentially druggable drivers have been identified [112]. Hence, the integration of expression data might prove beneficial in clarifying therapeutic decisions for patients with uncertain mutation-driven treatment options, which we illustrated with patients ACCC_CRC_05 and ACCC_CRC_06. Our integrated analysis suggested treatment options for these patients that were unknown at the time when the patients were in clinical treatment. Pathway and transcription factor network analysis suggested these two patients to be of the MSS subtype, even though the tumors had a very high TMB. We then identified five patients with similar characteristics in the TCGA cohort, all having very high TMB and carrying mutations in several genes involved in DNA damage response and DNA mismatch repair as well as in epigenetic writers. Furthermore, a particular frameshift mutation in *RPL22* has been causally related to missplicing of *MDM4* transcripts and loss of p53 signaling in colorectal cancer [113]. While this mutation was associated particularly with the MSI subtype, the p53 pathway is mostly

abrogated in the MSS subtype (compare Fig. 1, and Figs S2 and S3). Both ACCC tumors ACCC_CRC_05 and ACCC_CRC_06 contained this specific p.Lys15ArgfsTer5 mutation in the *RPL22* gene which, however, was not present in any of the five TCGA MSS patients. Thus, the two ACCC tumors as well as the five TCGA MSS tumors with high TMB showed characteristics of the MSS subtype, while other features suggested them belonging to the MSI subgroup. These tumors thus likely represent neither typical MSS nor MSI CRC.

The added value of integrating molecular modalities to foster precision medicine was initially introduced with the WINTHER clinical study [114] (NCT01856296). WINTHER was the first trial to describe the combination of matched genomic and transcriptomic information in metastatic tumors, showcased the significance of integrated molecular profiling into improved therapeutic recommendations, and expanded the portfolio of putative therapeutic interventions [114]. Likewise, the more recent MASTER/NCT clinical study has illustrated that the integration of WES/WGS with RNA sequencing can result in a clinical benefit and boost patient treatment stratification in patients with advanced cancer of rare cancer entities [63]. Our findings support these concepts, while we propose to enhance these concepts by integrating mutation ranking with immune deconvolution and pathway as well as transcription factor activity analysis, to suggest personalized treatment strategies as we did here for colorectal cancer. In the future, this will likely be complemented with proteomic information [115] and holds the promise for enhancing the biological rationale behind individualized therapeutic interventions even if it does not constitute a therapeutic protocol *per se*. Yet, it serves as a paradigm that supports and augments the concept of tailored medicine in oncology.

5. Conclusions

In this study, we elucidated the divergent genomic and transcriptomic landscapes of microsatellite stable (MSS) and microsatellite instable (MSI) colorectal cancers (CRCs) in a cohort from the Athens Comprehensive Cancer Center (ACCC) (https://www.accc.gr), by implementing an integrative bioinformatics workflow to compare sequence data from Athens, with large public multi-omics CRC datasets. Whole exome sequencing (WES) and RNA sequencing (RNAseq) of 28 CRC patients combined with integrated analysis of the two large CRC datasets of the TCGA and CPTAC cohorts revealed significant differences in mutational burden, variant types, and TF/pathway

activity levels between MSS and MSI subtypes. Our analysis not only corroborates previous findings, but also extends the biology of MSI-MSS CRC tumors and identifies potential vulnerabilities in individual tumors, advocating for the adoption of precision oncology approaches to improve clinical outcomes for CRC patients.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

GNZ clinically assessed or operated patients and coordinated sample acquisition. ST, GNZ, VK, SR, KoP, and AP collected and processed specimens. ST, SWo and SWi oversaw sequencing of the samples. E-IV, KV, GW, RK, KyP, and OP carried out the bioinformatics analyses and processed data. E-IV, KV, JB, OP, and SWi analyzed and interpreted data, and wrote the manuscript. E-IV, AD-S, GNZ, AP, and SWi conceived the study. All authors participated in discussions and interpretation of the data and results.

Peer review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/1878-0261.13814.

Data accessibility

All sequencing data supporting the findings of this study have been uploaded to the European Genome-

Phenome archive (http://www.ebi.ac.uk/ega/), and are available with controlled access of a DAC via the accession numbers EGAD00001011262 and EGAD 00001011263. All computational code, including R scripts, quarto documents and further information to reproduce the complete bioinformatics analyses, are available on Zenodo (https://zenodo.org/doi/10.5281/-zenodo.10959699) and Github (https://github.com/Jasonmbg/ACCC-CRC, https://github.com/Jasonmbg/TCGAPanCan-CRC and https://github.com/Jasonmbg/CPTAC-Colon) repositories for open access.

References

1 Casak SJ, Horiba MN, Yuan M, Cheng J, Lemery SJ, Shen YL, et al. FDA approval summary: tucatinib with trastuzumab for advanced unresectable or metastatic, chemotherapy refractory, HER2-positive RAS wild-type colorectal Cancer. *Clin Cancer Res*. 2023;**29**:4326–30. https://doi.org/10.1158/1078-0432.CCR-23-1041

2 Kopetz S, Grothey A, Yaeger R, Van Cutsem E, Desai J, Yoshino T, et al. Encorafenib, Binimetinib, and Cetuximab in BRAF V600E-mutated colorectal Cancer. *N Engl J Med*. 2019;**381**:1632–43. https://doi.org/10.1056/NEJMoa1908075

3 Janne PA, Riely GJ, Gadgeel SM, Heist RS, Ou SI, Pacheco JM, et al. Adagrasib in non-small-cell lung Cancer Harboring a KRAS(G12C) mutation. *N Engl J Med*. 2022;**387**:120–31. https://doi.org/10.1056/NEJMoa2204619

4 Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;**74**:229–63. https://doi.org/10.3322/caac.21834

5 Siegel RL, Wagle NS, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2023. *CA Cancer J Clin*. 2023;**73**:233–54. https://doi.org/10.3322/caac.21772

6 Scharpf RB, Balan A, Ricciuti B, Fiksel J, Cherry C, Wang C, et al. Genomic landscapes and hallmarks of mutant RAS in human cancers. *Cancer Res*. 2022;**82**:4058–78. https://doi.org/10.1158/0008-5472.CAN-22-1731

7 Dang HX, Krasnick BA, White BS, Grossman JG, Strand MS, Zhang J, et al. The clonal evolution of metastatic colorectal cancer. *Sci Adv*. 2020;**6**:eaay9691. https://doi.org/10.1126/sciadv.aay9691

8 Househam J, Heide T, Cresswell GD, Spiteri I, Kimberley C, Zapata L, et al. Phenotypic plasticity and genetic control in colorectal cancer evolution. *Nature*. 2022;**611**:744–53. https://doi.org/10.1038/s41586-022-05311-x

9 Loree JM, Pereira AAL, Lam M, Willauer AN, Raghav K, Dasari A, et al. Classifying colorectal Cancer by tumor location rather than sidedness highlights a continuum in mutation profiles and consensus molecular subtypes. *Clin Cancer Res*. 2018;**24**:1062–72. https://doi.org/10.1158/1078-0432.CCR-17-2484

10 Linares J, Sallent-Aragay A, Badia-Ramentol J, Recort-Bascuas A, Mendez A, Manero-Ruperez N, et al. Long-term platinum-based drug accumulation in cancer-associated fibroblasts promotes colorectal cancer progression and resistance to therapy. *Nat Commun*. 2023;**14**:746. https://doi.org/10.1038/s41467-023-36334-1

11 Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;**487**:330–7. https://doi.org/10.1038/nature11252

12 Mendelaar PAJ, Smid M, van Riet J, Angus L, Labots M, Steeghs N, et al. Whole genome sequencing of metastatic colorectal cancer reveals prior treatment effects and specific metastasis features. *Nat Commun*. 2021;**12**:574. https://doi.org/10.1038/s41467-020-20887-6

13 de la Chapelle A, Hampel H. Clinical relevance of microsatellite instability in colorectal cancer. *J Clin Oncol*. 2010;**28**:3380–7. https://doi.org/10.1200/JCO.2009.27.0652

14 Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, et al. A National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res*. 1998;**58**:5248–57.

15 Lochhead P, Kuchiba A, Imamura Y, Liao X, Yamauchi M, Nishihara R, et al. Microsatellite instability and BRAF mutation testing in colorectal cancer prognostication. *J Natl Cancer Inst*. 2013;**105**:1151–6. https://doi.org/10.1093/jnci/djt173

16 Poynter JN, Siegmund KD, Weisenberger DJ, Long TI, Thibodeau SN, Lindor N, et al. Molecular characterization of MSI-H colorectal cancer by MLHI promoter methylation, immunohistochemistry, and mismatch repair germline mutation screening. *Cancer Epidemiol Biomarkers Prev*. 2008;**17**:3208–15. https://doi.org/10.1158/1055-9965.EPI-08-0512

17 Lynch HT, Snyder CL, Shaw TG, Heinen CD, Hitchins MP. Milestones of Lynch syndrome: 1895–2015. *Nat Rev Cancer*. 2015;**15**:181–94. https://doi.org/10.1038/nrc3878

18 Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med*. 2015;**372**:2509–20. https://doi.org/10.1056/NEJMoa1500596

19 Innocenti F, Mu W, Qu X, Ou FS, Kabbarah O, Blanke CD, et al. DNA mutational profiling in patients with colorectal Cancer treated with standard of care reveals differences in outcome and racial distribution of mutations. *J Clin Oncol*. 2023;**42**: JCO2300825. https://doi.org/10.1200/JCO.23.00825

20 De Smedt L, Lemahieu J, Palmans S, Govaere O, Tousseyn T, Van Cutsem E, et al. Microsatellite instable vs stable colon carcinomas: analysis of tumour heterogeneity, inflammation and angiogenesis. *Br J Cancer*. 2015;**113**:500–9. https://doi.org/10.1038/bjc.2015.213

21 Valdeolivas A, Amberg B, Giroud N, Richardson M, Galvez EJC, Badillo S, et al. Profiling the heterogeneity of colorectal cancer consensus molecular subtypes using spatial transcriptomics. *NPJ Precis Oncol*. 2024;**8**:10. https://doi.org/10.1038/s41698-023-00488-4

22 Reisinger E, Genthner L, Kerssemakers J, Kensche P, Borufka S, Jugold A, et al. OTP: an automatized system for managing and processing NGS data. *J Biotechnol*. 2017;**261**:53–62. https://doi.org/10.1016/j.jbiotec.2017.08.006

23 Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, et al. The repertoire of mutational signatures in human cancer. *Nature*. 2020;**578**:94–101. https://doi.org/10.1038/s41586-020-1943-3

24 Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;**38**:e164. https://doi.org/10.1093/nar/gkq603

25 Pagel KA, Kim R, Moad K, Busby B, Zheng L, Tokheim C, et al. Integrated informatics analysis of Cancer-related variants. *JCO Clin Cancer Inform*. 2020;**4**:310–7. https://doi.org/10.1200/CCI.19.00132

26 Wagner AH, Walsh B, Mayfield G, Tamborero D, Sonkin D, Krysiak K, et al. A harmonized meta-knowledgebase of clinical interpretations of somatic genomic variants in cancer. *Nat Genet*. 2020;**52**:448–57. https://doi.org/10.1038/s41588-020-0603-8

27 Froyen G, Le Mercier M, Lierman E, Vandepoele K, Nollet F, Boone E, et al. Standardization of somatic variant classifications in solid and haematological tumours by a two-level approach of biological and clinical classes: an initiative of the Belgian ComPerMed expert panel. *Cancers (Basel)*. 2019;**11**:2030. https://doi.org/10.3390/cancers11122030

28 Morales J, Pujar S, Loveland JE, Astashyn A, Bennett R, Berry A, et al. A joint NCBI and EMBL-EBI transcript set for clinical genomics and research. *Nature*. 2022;**604**:310–5. https://doi.org/10.1038/s41586-022-04558-8

29 Cunningham F, Moore B, Ruiz-Schultz N, Ritchie GR, Eilbeck K. Improving the sequence ontology terminology for genomic variant annotation. *J Biomed Semantics*. 2015;**6**:32. https://doi.org/10.1186/s13326-015-0030-4

30 Martinez-Jimenez F, Muinos F, Sentis I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, et al. A compendium of mutational cancer driver genes. *Nat Rev Cancer*. 2020;**20**:555–72. https://doi.org/10.1038/s41568-020-0290-x

31 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**:e1004873. https://doi.org/10.1371/journal.pcbi.1004873

32 Kleinheinz K, Bludau I, Hübschmann D, Heinold M, Kensche P, Gu Z, et al. ACEseq – allele specific copy number estimation from whole genome sequencing. *bioRxiv*. 2017 https://doi.org/10.1101/210807

33 Olshen AB, Bengtsson H, Neuvial P, Spellman PT, Olshen RA, Seshan VE. Parent-specific copy number in paired tumor-normal studies using circular binary segmentation. *Bioinformatics*. 2011;**27**:2038–46. https://doi.org/10.1093/bioinformatics/btr329

34 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;**29**:15–21. https://doi.org/10.1093/bioinformatics/bts635

35 DeLuca DS, Levin JZ, Sivachenko A, Fennell T, Nazaire MD, Williams C, et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics*. 2012;**28**:1530–2. https://doi.org/10.1093/bioinformatics/bts196

36 Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;**30**:923–30. https://doi.org/10.1093/bioinformatics/btt656

37 Liao Y, Smyth GK, Shi W. The subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res*. 2013;**41**:e108. https://doi.org/10.1093/nar/gkt214

38 Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;**26**:139–40. https://doi.org/10.1093/bioinformatics/btp616

39 Sturm G, Finotello F, Petitprez F, Zhang JD, Baumbach J, Fridman WH, et al. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics*. 2019;**35**:i436–45. https://doi.org/10.1093/bioinformatics/btz363

40 Vasaikar S, Huang C, Wang X, Petyuk VA, Savage SR, Wen B, et al. Proteogenomic analysis of human colon Cancer reveals new therapeutic opportunities. *Cell*. 2019;**177**:1035–1049 e1019. https://doi.org/10.1016/j.cell.2019.03.030

41 Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, et al. Proteogenomic characterization of human colon and rectal cancer. *Nature*. 2014;**513**:382–7. https://doi.org/10.1038/nature13438

42 Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;**28**:1747–56. https://doi.org/10.1101/gr.239244.118

43 Kautto EA, Bonneville R, Miya J, Yu L, Krook MA, Reeser JW, et al. Performance evaluation for rapid detection of pan-cancer microsatellite instability with MANTIS. *Oncotarget*. 2017;**8**:7452–63. https://doi.org/10.18632/oncotarget.13918

44 Li L, Feng Q, Wang X. PreMSIm: an R package for predicting microsatellite instability from the expression profiling of a gene panel in cancer. *Comput Struct Biotechnol J*. 2020;**18**:668–75. https://doi.org/10.1016/j.csbj.2020.03.007

45 Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;**16**:284–7. https://doi.org/10.1089/omi.2011.0118

46 Yu G, He QY. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol Biosyst*. 2016;**12**:477–9. https://doi.org/10.1039/c5mb00663e

47 Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, et al. The reactome pathway knowledgebase 2022. *Nucleic Acids Res*. 2022;**50**:D687–92. https://doi.org/10.1093/nar/gkab1028

48 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;**102**:15545–50. https://doi.org/10.1073/pnas.0506580102

49 Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature*. 2009;**462**:108–12. https://doi.org/10.1038/nature08460

50 Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;**1**:417–25. https://doi.org/10.1016/j.cels.2015.12.004

51 Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res*. 2023;**51**:D638–46. https://doi.org/10.1093/nar/gkac1000

52 Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies.

Nucleic Acids Res. 2015;**43**:e47. https://doi.org/10.1093/nar/gkv007

53 Schubert M, Klinger B, Klunemann M, Sieber A, Uhlitz F, Sauer S, et al. Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat Commun*. 2018;**9**:20. https://doi.org/10.1038/s41467-017-02391-6

54 Holland CH, Tanevski J, Perales-Paton J, Gleixner J, Kumar MP, Mereu E, et al. Robustness and applicability of transcription factor and pathway analysis tools on single-cell RNA-seq data. *Genome Biol*. 2020;**21**:36. https://doi.org/10.1186/s13059-020-1949-z

55 Badia IMP, Velez Santiago J, Braunger J, Geiss C, Dimitrov D, Muller-Dott S, et al. decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinform Adv*. 2022;**2**:vbac016. https://doi.org/10.1093/bioadv/vbac016

56 Wickham H. ggplot2: elegant graphics for data analysis. New York, NY: Springer; 2009.

57 Krug K, Mertins P, Zhang B, Hornbeck P, Raju R, Ahmad R, et al. A curated resource for Phosphosite-specific signature analysis. *Mol Cell Proteomics*. 2019;**18**:576–93. https://doi.org/10.1074/mcp.TIR118.000943

58 Turei D, Valdeolivas A, Gul L, Palacio-Escat N, Klein M, Ivanova O, et al. Integrated intra- and intercellular signaling knowledge for multicellular omics analysis. *Mol Syst Biol*. 2021;**17**:e9923. https://doi.org/10.15252/msb.20209923

59 Bachman JA, Sorger PK, Gyori BM. Assembling a corpus of phosphoproteomic annotations using ProtMapper to normalize site information from databases and text mining. *bioRxiv*. 2022 https://doi.org/10.1101/822668

60 Lê S, Josse J, Husson F. FactoMineR: AnRPackage for multivariate analysis. *J Stat Softw*. 2008;**25**:1–18. https://doi.org/10.18637/jss.v025.i01

61 Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;**32**:2847–9. https://doi.org/10.1093/bioinformatics/btw313

62 Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, et al. OncoKB: a precision oncology Knowledge Base. *JCO Precis Oncol*. 2017;**1**:1–16. https://doi.org/10.1200/PO.17.00011

63 Horak P, Heining C, Kreutzfeldt S, Hutter B, Mock A, Hullein J, et al. Comprehensive genomic and transcriptomic analysis for guiding therapeutic decisions in patients with rare cancers. *Cancer Discov*. 2021;**11**:2780–95. https://doi.org/10.1158/2159-8290.CD-21-0126

64 Edwards NJ, Oberti M, Thangudu RR, Cai S, McGarvey PB, Jacob S, et al. The CPTAC data portal: a resource for Cancer proteomics research. *J*

Proteome Res. 2015;**14**:2707–13. https://doi.org/10.1021/pr501254j

65 Zampino G, Pantaleoni F, Carta C, Cobellis G, Vasta I, Neri C, et al. Diversity, parental germline origin, and phenotypic spectrum of de novo HRAS missense changes in Costello syndrome. *Hum Mutat.* 2007;**28**:265–72. https://doi.org/10.1002/humu.20431

66 Hu J, Stites EC, Yu H, Germino EA, Meharena HS, Stork PJS, et al. Allosteric activation of functionally asymmetric RAF kinase dimers. *Cell.* 2013;**154**:1036–46. https://doi.org/10.1016/j.cell.2013.07.046

67 Choi J, Landrette SF, Wang T, Evans P, Bacchiocchi A, Bjornson R, et al. Identification of PLX4032-resistance mechanisms and implications for novel RAF inhibitors. *Pigment Cell Melanoma Res.* 2014;**27**:253–62. https://doi.org/10.1111/pcmr.12197

68 de Gooyer PGM, Verschoor YL, van den Dungen LDW, Balduzzi S, Marsman HA, Geukes Foppen MH, et al. Neoadjuvant nivolumab and relatlimab in locally advanced MMR-deficient colon cancer: a phase 2 trial. *Nat Med.* 2024;**30**:3284–90. https://doi.org/10.1038/s41591-024-03250-w

69 Dienstmann R, Vermeulen L, Guinney J, Kopetz S, Tejpar S, Tabernero J. Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nat Rev Cancer.* 2017;**17**:79–92. https://doi.org/10.1038/nrc.2016.126

70 Smeets D, Miller IS, O'Connor DP, Das S, Moran B, Boeckx B, et al. Copy number load predicts outcome of metastatic colorectal cancer patients receiving bevacizumab combination therapy. *Nat Commun.* 2018;**9**:4112. https://doi.org/10.1038/s41467-018-06567-6

71 Zhang D, Ni QQ, Liang QY, He LL, Qiu BW, Zhang LJ, et al. ASCL2 induces an immune excluded microenvironment by activating cancer-associated fibroblasts in microsatellite stable colorectal cancer. *Oncogene.* 2023;**42**:2841–53. https://doi.org/10.1038/s41388-023-02806-3

72 Nguyen TT, Ramsay L, Ahanfeshar-Adams M, Lajoie M, Schadendorf D, Alain T, et al. Mutations in the IFNgamma-JAK-STAT pathway causing resistance to immune checkpoint inhibitors in melanoma increase sensitivity to oncolytic virus treatment. *Clin Cancer Res.* 2021;**27**:3432–42. https://doi.org/10.1158/1078-0432.CCR-20-3365

73 Coles GL, Cristea S, Webber JT, Levin RS, Moss SM, He A, et al. Unbiased proteomic profiling uncovers a targetable GNAS/PKA/PP2A Axis in small cell lung Cancer stem cells. *Cancer Cell.* 2020;**38**:129–143 e127. https://doi.org/10.1016/j.ccell.2020.05.003

74 Ito Y, Bae SC, Chuang LS. The RUNX family: developmental regulators in cancer. *Nat Rev Cancer.* 2015;**15**:81–95. https://doi.org/10.1038/nrc3877

75 Mevel R, Draper JE, Lie ALM, Kouskoff V, Lacaud G. RUNX transcription factors: orchestrators of development. *Development.* 2019;**146**:dev148296. https://doi.org/10.1242/dev.148296

76 Xu H, Liu L, Li W, Zou D, Yu J, Wang L, et al. Transcription factors in colorectal cancer: molecular mechanism and therapeutic implications. *Oncogene.* 2021;**40**:1555–69. https://doi.org/10.1038/s41388-020-01587-3

77 Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med.* 2015;**21**:1350–6. https://doi.org/10.1038/nm.3967

78 Nitsche U, Zimmermann A, Spath C, Muller T, Maak M, Schuster T, et al. Mucinous and signet-ring cell colorectal cancers differ from classical adenocarcinomas in tumor biology and prognosis. *Ann Surg.* 2013;**258**:775–82; discussion 782–773. https://doi.org/10.1097/SLA.0b013e3182a69f7e

79 Linnekamp JF, Hooff SRV, Prasetyanti PR, Kandimalla R, Buikhuisen JY, Fessler E, et al. Consensus molecular subtypes of colorectal cancer are recapitulated in in vitro and in vivo models. *Cell Death Differ.* 2018;**25**:616–33. https://doi.org/10.1038/s41418-017-0011-5

80 Oshima S, Nakamura T, Namiki S, Okada E, Tsuchiya K, Okamoto R, et al. Interferon regulatory factor 1 (IRF-1) and IRF-2 distinctively up-regulate gene expression and production of interleukin-7 in human intestinal epithelial cells. *Mol Cell Biol.* 2004;**24**:6298–310. https://doi.org/10.1128/MCB.24.14.6298-6310.2004

81 Liao W, Overman MJ, Boutin AT, Shang X, Zhao D, Dey P, et al. KRAS-IRF2 Axis drives immune suppression and immune therapy resistance in colorectal Cancer. *Cancer Cell.* 2019;**35**:559–572 e557. https://doi.org/10.1016/j.ccell.2019.02.008

82 Prahallad A, Sun C, Huang S, Di Nicolantonio F, Salazar R, Zecchin D, et al. Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature.* 2012;**483**:100–3. https://doi.org/10.1038/nature10868

83 Troiani T, Napolitano S, Vitagliano D, Morgillo F, Capasso A, Sforza V, et al. Primary and acquired resistance of colorectal cancer cells to anti-EGFR antibodies converge on MEK/ERK pathway activation and can be overcome by combined MEK/EGFR inhibition. *Clin Cancer Res.* 2014;**20**:3775–86. https://doi.org/10.1158/1078-0432.CCR-13-2181

84 Kim D, Herdeis L, Rudolph D, Zhao Y, Bottcher J, Vides A, et al. Pan-KRAS inhibitor disables oncogenic signalling and tumour growth. *Nature.* 2023;**619**:160–6. https://doi.org/10.1038/s41586-023-06123-3

85 Pahuja KB, Nguyen TT, Jaiswal BS, Prabhash K, Thaker TM, Senger K, et al. Actionable activating oncogenic ERBB2/HER2 transmembrane and Juxtamembrane domain mutations. *Cancer Cell.*

2018;**34**:792–806 e795. https://doi.org/10.1016/j.ccell.2018.09.010

86 Wang C, Jette N, Moussienko D, Bebb DG, Lees-Miller SP. ATM-deficient colorectal Cancer cells are sensitive to the PARP inhibitor Olaparib. *Transl Oncol*. 2017;**10**:190–6. https://doi.org/10.1016/j.tranon.2017.01.007

87 Behl D, Rothe M, Mangat PK, Garrett-Mayer E, Farrington LC, Crysler OV, et al. Olaparib (O) in patients (pts) with colorectal cancer (CRC) with ATM mutation (mut): results from the targeted agent and profiling utilization Registry (TAPUR) study. *J Clin Oncol*. 2023;**41**:122. https://doi.org/10.1200/JCO.2023.41.4_suppl.122

88 Neshat MS, Mellinghoff IK, Tran C, Stiles B, Thomas G, Petersen R, et al. Enhanced sensitivity of PTEN-deficient tumors to inhibition of FRAP/mTOR. *Proc Natl Acad Sci U S A*. 2001;**98**:10314–9. https://doi.org/10.1073/pnas.171076798

89 Ma L, Teruya-Feldstein J, Bonner P, Bernardi R, Franz DN, Witte D, et al. Identification of S664 TSC2 phosphorylation as a marker for extracellular signal-regulated kinase mediated mTOR activation in tuberous sclerosis and human cancer. *Cancer Res*. 2007;**67**:7106–12. https://doi.org/10.1158/0008-5472.CAN-06-4798

90 Fritsche-Guenther R, Zasada C, Mastrobuoni G, Royla N, Rainer R, Rossner F, et al. Alterations of mTOR signaling impact metabolic stress resistance in colorectal carcinomas with BRAF and KRAS mutations. *Sci Rep*. 2018;**8**:9204. https://doi.org/10.1038/s41598-018-27394-1

91 Jan YH, Tan KT, Chen SJ, Yip TTC, Lu CT, Lam AK. Comprehensive assessment of actionable genomic alterations in primary colorectal carcinoma using targeted next-generation sequencing. *Br J Cancer*. 2022;**127**:1304–11. https://doi.org/10.1038/s41416-022-01913-4

92 Westcott PMK, Muyas F, Hauck H, Smith OC, Sacks NJ, Ely ZA, et al. Mismatch repair deficiency is not sufficient to elicit tumor immunogenicity. *Nat Genet*. 2023;**55**:1686–95. https://doi.org/10.1038/s41588-023-01499-4

93 Ballhausen A, Przybilla MJ, Jendrusch M, Haupt S, Pfaffendorf E, Seidler F, et al. The shared frameshift mutation landscape of microsatellite-unstable cancers suggests immunoediting during tumor evolution. *Nat Commun*. 2020;**11**:4740. https://doi.org/10.1038/s41467-020-18514-5

94 Becker WR, Nevins SA, Chen DC, Chiu R, Horning AM, Guha TK, et al. Single-cell analyses define a continuum of cell state and composition changes in the malignant transformation of polyps to colorectal cancer. *Nat Genet*. 2022;**54**:985–95. https://doi.org/10.1038/s41588-022-01088-x

95 Li Q, Lai Q, He C, Fang Y, Yan Q, Zhang Y, et al. RUNX1 promotes tumour metastasis by activating the Wnt/beta-catenin signalling pathway and EMT in colorectal cancer. *J Exp Clin Cancer Res*. 2019;**38**:334. https://doi.org/10.1186/s13046-019-1330-9

96 Pant S, Schuler M, Iyer G, Witt O, Doi T, Qin S, et al. Erdafitinib in patients with advanced solid tumours with FGFR alterations (RAGNAR): an international, single-arm, phase 2 study. *Lancet Oncol*. 2023;**24**:925–35. https://doi.org/10.1016/S1470-2045(23)00275-9

97 Jung B, Staudacher JJ, Beauchamp D. Transforming growth factor beta superfamily Signaling in development of colorectal Cancer. *Gastroenterology*. 2017;**152**:36–52. https://doi.org/10.1053/j.gastro.2016.10.015

98 Xue Z, Vis DJ, Bruna A, Sustic T, van Wageningen S, Batra AS, et al. MAP3K1 and MAP2K4 mutations are associated with sensitivity to MEK inhibitors in multiple cancer models. *Cell Res*. 2018;**28**:719–29. https://doi.org/10.1038/s41422-018-0044-4

99 Ried T, Meijer GA, Harrison DJ, Grech G, Franch-Exposito S, Briffa R, et al. The landscape of genomic copy number alterations in colorectal cancer and their consequences on gene expression levels and disease outcome. *Mol Aspects Med*. 2019;**69**:48–61. https://doi.org/10.1016/j.mam.2019.07.007

100 Lord CJ, Ashworth A. BRCAness revisited. *Nat Rev Cancer*. 2016;**16**:110–20. https://doi.org/10.1038/nrc.2015.21

101 Mestrallet G, Brown M, Bozkus CC, Bhardwaj N. Immune escape and resistance to immunotherapy in mismatch repair deficient tumors. *Front Immunol*. 2023;**14**:1210164. https://doi.org/10.3389/fimmu.2023.1210164

102 Guo Y, Lu X, Chen Y, Clark G, Trent J, Cuatrecasas M, et al. Opposing roles of ZEB1 in the cytoplasm and nucleus control cytoskeletal assembly and YAP1 activity. *Cell Rep*. 2022;**41**:111452. https://doi.org/10.1016/j.celrep.2022.111452

103 Huang Y, Hong W, Wei X. The molecular mechanisms and therapeutic strategies of EMT in tumor progression and metastasis. *J Hematol Oncol*. 2022;**15**:129. https://doi.org/10.1186/s13045-022-01347-8

104 Lal N, White BS, Goussous G, Pickles O, Mason MJ, Beggs AD, et al. KRAS mutation and consensus molecular subtypes 2 and 3 are independently associated with reduced immune infiltration and reactivity in colorectal Cancer. *Clin Cancer Res*. 2018;**24**:224–33. https://doi.org/10.1158/1078-0432.CCR-17-1090

105 Garralda E, Beaulieu ME, Moreno V, Casacuberta-Serra S, Martinez-Martin S, Foradada L, et al. MYC targeting by OMO-103 in solid tumors: a phase 1 trial. *Nat Med*. 2024;**30**:762–71. https://doi.org/10.1038/s41591-024-02805-1

106 Li X, Kaur N, Albahrani M, Karpf AR, Black AR, Black JD. Crosstalk between protein kinase C alpha and transforming growth factor beta signaling mediated by Runx2 in intestinal epithelial cells. *J Biol Chem*. 2023;**299**:103017. https://doi.org/10.1016/j.jbc.2023.103017

107 Yan X, Han D, Chen Z, Han C, Dong W, Han L, et al. RUNX2 interacts with BRG1 to target CD44 for promoting invasion and migration of colorectal cancer cells. *Cancer Cell Int*. 2020;**20**:505. https://doi.org/10.1186/s12935-020-01544-w

108 Tuo Z, Zhang Y, Wang X, Dai S, Liu K, Xia D, et al. RUNX1 is a promising prognostic biomarker and related to immune infiltrates of cancer-associated fibroblasts in human cancers. *BMC Cancer*. 2022;**22**:523. https://doi.org/10.1186/s12885-022-09632-y

109 Heide T, Househam J, Cresswell GD, Spiteri I, Lynn C, Mossner M, et al. The co-evolution of the genome and epigenome in colorectal cancer. *Nature*. 2022;**611**:733–43. https://doi.org/10.1038/s41586-022-05202-1

110 Joanito I, Wirapati P, Zhao N, Nawaz Z, Yeo G, Lee F, et al. Single-cell and bulk transcriptome sequencing identifies two epithelial tumor cell states and refines the consensus molecular classification of colorectal cancer. *Nat Genet*. 2022;**54**:963–75. https://doi.org/10.1038/s41588-022-01100-4

111 Khaliq AM, Erdogan C, Kurt Z, Turgut SS, Grunvald MW, Rand T, et al. Refining colorectal cancer classification and clinical stratification through a single-cell atlas. *Genome Biol*. 2022;**23**:113. https://doi.org/10.1186/s13059-022-02677-z

112 Mateo L, Duran-Frigola M, Gris-Oliver A, Palafox M, Scaltriti M, Razavi P, et al. Personalized cancer therapy prioritization based on driver alteration co-occurrence patterns. *Genome Med*. 2020;**12**:78. https://doi.org/10.1186/s13073-020-00774-x

113 Weinstein HNW, Hu K, Fish L, Chen YA, Allegakoen P, Pham JH, et al. RPL22 is a tumor suppressor in MSI-high cancers and a splicing regulator of MDM4. *Cell Rep*. 2024;**43**:114622. https://doi.org/10.1016/j.celrep.2024.114622

114 Rodon J, Soria JC, Berger R, Miller WH, Rubin E, Kugel A, et al. Genomic and transcriptomic profiling expands precision cancer medicine: the WINTHER trial. *Nat Med*. 2019;**25**:751–8. https://doi.org/10.1038/s41591-019-0424-4

115 Wahjudi LW, Bernhardt S, Abnaof K, Horak P, Kreutzfeldt S, Heining C, et al. Integrating proteomics into precision oncology. *Int J Cancer*. 2021;**148**:1438–51. https://doi.org/10.1002/ijc.33301

116 Manders F, Brandsma AM, de Kanter J, Verheul M, Oka R, van Roosmalen MJ, et al. MutationalPatterns: the one stop shop for the analysis of mutational processes. *BMC Genomics*. 2022;**23**:134. https://doi.org/10.1186/s12864-022-08357-3

117 Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, et al. GO::TermFinder--open source software for accessing gene ontology information and finding significantly enriched gene ontology terms associated with a list of genes. *Bioinformatics*. 2004;**20**:3710–5. https://doi.org/10.1093/bioinformatics/bth456

118 Horak P, Leichsenring J, Goldschmid H, Kreutzfeldt S, Kazdal D, Teleanu V, et al. Assigning evidence to actionability: an introduction to variant interpretation in precision cancer medicine. *Genes Chromosomes Cancer*. 2022;**61**:303–13. https://doi.org/10.1002/gcc.22987

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Relative contribution of COSMIC mutational signatures in each patient of the Athens Comprehensive Cancer Center Colorectal Cancer (ACCC-CRC) cohort.

Fig. S2. Oncoplot of 54 cancer driver genes across 434 colorectal adenocarcinoma samples from The Cancer Genome Atlas Network (TCGA) PanCancer Atlas cohort (392 microsatellite stable-MSS and 42 microsatellite instable-MSI cases).

Fig. S3. Oncoplot of 54 cancer driver genes across 79 colorectal adenocarcinoma samples from the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC) cohort (65 microsatellite stable-MSS and 14 microsatellite instable-MSI cases).

Fig. S4. K-means clustering of STRING protein–protein interaction network of high mutational frequency genes in microsatellite instable (MSI) vs. microsatellite stable (MSS) colorectal cancer.

Fig. S5. Kinase and Pathway activities, and GNAS alterations in microsatellite stable (MSS) and microsatellite instable (MSI) colorectal cancer (CRC) tumors of the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC) cohort.

Fig. S6. Clustering of the colorectal cancer (CRC) patients of The Cancer Genome Atlas Network (TCGA) based on transcriptomic analysis.

Fig. S7. Clustering of the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium Colon Adenocarcinoma (CPTAC-COAD) dataset based on transcriptomic profiling.

Table S1. Patient characteristics of the Athens Comprehensive Cancer Center Colorectal Cancer (ACCC-CRC) patient cohort.

Table S2. Microsatellite instability (MSI)-status (MSI MANTIS [43]) and Tumor Mutational Burden (TMB)

values as provided in The Cancer Genome Atlas Network (TCGA) dataset, and PreMSIm [44] predictions for 434 patients of the TCGA PanCancer Atlas COAD/READ cohort [11].

Table S3. Microsatellite instability (MSI)-status and Tumor Mutational Burden (TMB) values as provided in the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC) dataset, and PreMSIm [44] predictions for 79 patients of the CPTAC colorectal cancer (CRC) cohort [64].

Table S4. Tumor mutational burden, total numbers of somatic mutations, numbers of SVRACAS scored mutations, microsatellite instability (MSI)-status prediction and probability for 28 patients of the Athens Comprehensive Cancer Center Colorectal Cancer (ACCC-CRC) cohort.

Table S5. List of somatic mutations in the Athens Comprehensive Cancer Center Colorectal Cancer (ACCC-CRC) patients, with variant annotation information from OpenCRAVAT/OakVar, along with the SVRACAS scoring estimates.

Table S6. List of somatic mutations in The Cancer Genome Atlas Network (TCGA) colorectal cancer (CRC) patients, with variant annotation information from OpenCRAVAT/OakVar, along with the SVRACAS scoring estimates.

Table S7. List of somatic mutations in the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC) colorectal cancer (CRC) patients, with variant annotation information from OpenCRAVAT/OakVar, along with the SVRACAS scoring estimates.