Molecular and functional profiling unravels targetable vulnerabilities in colorectal cancer

Efstathios-Iason Vlachavas[1,#,*], Konstantinos Voutetakis[2,#], 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,*], and Stefan Wiemann[1,*]

(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) University Hospital of North Midlands, Surgical Department, 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

24 (11) Department of Medicine II, University Medical Center Mannheim, Medical Faculty
25 Mannheim, Heidelberg University, Mannheim, Germany
26 (12) DKFZ-Hector Cancer Institute at University Medical Center Mannheim, Germany
27 # equal contribution

28 * corresponding authors

29 Abstract

30 While colorectal cancer (CRC) patients with microsatellite instability (MSI) respond well to
31 immunotherapy those with microsatellite stable (MSS) tumors rely on conventional chemotherapy,
32 often with poor outcomes. Both types frequently carry mutations in *KRAS* or *BRAF* proto-oncogenes,
33 rendering them more resistant to treatment. New therapeutic biomarkers and treatments remain a
34 clinical need, especially for MSS tumors. We performed whole exome and RNA-Sequencing from 28
35 tumors of the Athens Comprehensive Cancer Center CRC cohort, and molecularly characterized
36 colorectal cancer patients based on their MSI status, SNVs/CNAs, and pathway/transcription factor
37 activities at the individual patient level. Variants were classified using a new computational score for
38 integrative cancer variant annotation and prioritization. Complementing this molecular data with
39 public multi-omics datasets, we identified activation of transforming growth factor beta (TGFβ)
40 signaling to be stronger activated in the MSS patients whereas JAK-STAT and MAPK molecular
41 cascades were activated specifically in MSI. We unraveled mechanisms consistently perturbed in the
42 transcriptional and mutational circuits and identified RUNX transcription factors as putative novel
43 targets. Assessing the immunogenicity of CRC tumors in the context of RAS/RAF mutations and
44 MSI/MSS status revealed a critical impact *KRAS* mutations have on immunogenicity particularly in the
45 MSS patient subgroup, with implications for diagnosis and treatment.

46 Introduction

47 Colorectal cancer (CRC) is a highly heterogeneous disease associated with different molecular
48 background and clinical management. Improved screening strategies, surgical techniques and
49 advanced treatment protocols have led to better clinical outcomes of CRC patients in the last
50 decades [1, 2]. The recent development of covalent inhibitors specifically targeting the KRAS G12C
51 variant have indicated significant progress in targeting even previously 'undruggable' onco-proteins [3].
52 Despite these advances, CRC has remained the second and third leading cause of cancer-related
53 death in the US and worldwide, respectively [4, 5]. Furthermore, recent epidemiological data has
54 indicated a concerning trend of increased CRC incidence particularly in younger individuals below the
55 age of 55, coupled with more advanced disease stages [4]. The response to treatment is highly variable
56 and not durable in the majority of cases. This dismal clinical perspective is impacted by various
57 factors, including tissue tropism effects that allele-specific *RAS* mutants have in combination with
58 other actionable alterations [6], clonal heterogeneity [7], phenotypic plasticity [8], location of tumor [9], and
59 the tumor microenvironment [10]. Collectively, these findings highlight the need for novel therapies
60 focused specifically on the personalized treatment of CRC. The advent of next-generation sequencing
61 has led to substantial progress in molecular characterization of the disease. Alterations in a number
62 of genes with critical roles in CRC have been identified, such as in *APC, BRAF, EGFR, NRAS, KRAS* and
63 *PIK3CA* [11, 12]. The resulting aberrant activation of signaling pathways may be exploited in precision
64 therapies. Recent accomplishments include combinatorial strategies targeting co-activated pathways
65 in various types of unresectable CRC such as BRAF V600E mutated [1] and *HER2*-positive carcinomas [2].

66 The microsatellite instability (MSI) status is among the most informative parameters determining
67 treatment decision towards personalized medicine in CRC [13], and in other tumor entities [14]. According
68 to the MSI status, CRC is categorized into three types: high microsatellite instable (MSI-H), low
69 microsatellite instable (MSI-L) and microsatellite stabile (MSS) [13]. MSI-L and MSS tumors are
70 commonly grouped into one type [15]. Accordingly, we use MSS for MSI-L/MSS and MSI for MSI-H
71 tumors throughout. MSI-status accounts for ~15% of CRC cases and is associated with sporadic or
72 hereditary disease etiologies. While the former is associated with *BRAF* mutations [16] and *MLH1*

73 promoter methylation [17], the latter has been associated with mutations mostly in mismatch repair
74 genes *MLH1*, *MSH2*, *MSH6* as well as *PMS2* and is classified as Lynch syndrome [18]. MSI tumors have a
75 high mutational burden, resulting in the expression of neo-antigens that may be recognized by the
76 immune system. Differences in the immune microenvironment thus distinguish MSI and MSS CRCs,
77 and have immediate therapeutic impact: MSI CRC tumors mostly respond to immunotherapy, which
78 has revolutionized clinical practice and outcome of respective patients [19].

79 In contrast, MSS CRC patients commonly do not benefit from immunotherapies and have mostly
80 limited targeted treatment options, wherein the most informative molecular distinction involves the
81 *RAS* mutation status [20]. Despite the unmet therapeutic demand particularly for MSS patients, only
82 few studies have investigated the multi-layered differences in MSI/MSS tumors, their downstream
83 signaling effects, and their interrelation with *RAS*/*RAF* mutations in regard to immunogenicity and
84 pathway activation [21, 22]. Hence, by leveraging a prospective patient cohort of the Athens
85 Comprehensive Cancer Center (ACCC - https://www.accc.gr/index.html) and two public cancer
86 genomics datasets, we aspired to elucidate the molecular complexity of microsatellite (in)stability
87 status in CRC and to uncover biological mechanisms that might be exploited to enhance therapeutic
88 options particularly for patients with the MSS type.

89 Results

90 The mutational landscapes of MSS and MSI CRCs

91 To elucidate the landscape of somatic mutations in CRC, whole exome sequencing (WES) as well as
92 RNA sequencing (RNAseq) was performed in cancerous as well as in adjacent normal tissues from 28
93 patients of a prospective patient cohort of the Athens Comprehensive Cancer Center (ACCC)
94 (Supplementary Table 1). Since microsatellite instability is one of the major clinical factors
95 determining therapeutic management and prognosis of CRC patients, we first wanted to infer the
96 microsatellite stability status of every patient using RNAseq data. To this end, we initially validated
97 the robustness of PreMSIm [23] and predicted the MSI status in 434 patients of the TCGA PanCan

98 COAD/READ [11] as well as in 79 patients of the CPTAC [24] cohorts. PreMSIm estimates were highly
99 concordant with MSI-MANTIS predictions from TCGA (concordance rate: 0.988, Supplementary Table
100 2), and with the annotated MSI-status of patients in CPTAC (concordance rate: 0.987, Supplementary
101 Table 3). We then characterized the ACCC CRC samples applying PreMSIm and inferred five tumors as
102 MSI and 23 as MSS (Supplementary Table 4).

103 The exomes of ACCC CRC samples were sequenced at an average coverage of >300x (range: 195 to
104 375). In total, 25,913 unique protein-altering somatic variants were identified in 11,492 genes,
105 including 21,917 missense, 1495 stop_gained, 423 splice site, 1871 frameshift and 162 inframe InDel
106 mutations (Supplementary Table 5). The number of somatic variants varied greatly among ACCC
107 patients as the tumor mutational burden (TMB) ranged from 0.028 to 375.994 protein-altering
108 mutations per Megabase (Supplementary Table 4). The median was 108 nonsynonymous SNVs in
109 MSS tumors while that in MSI tumors was 1257. Five patients carried hot-spot mutations in the G12
110 codon of KRAS [G12D (n=3), G12V (n=1), G12S (n=1)], another five patients carried hot-spot
111 mutations in BRAF [V600E (n=4), K601E (n=1)], while concomitant mutations in KRAS [A146V] and
112 BRAF [D594N] were observed in one ACCC-patient. Another patient carried an HRAS [A146T]
113 mutation co-occurring with two non-hotspot mutations in BRAF [L505F and L577I]. The HRAS [A146T]
114 substitution has been found as a germline variant in Costello's syndrome [25] and as a somatic
115 mutation in a few cases of melanoma and other skin cancers (https://www.cbioportal.org, GENIE
116 Cohort v14.1-public, accessed on 2023/12/04). BRAF [L505F] is a class II mutation [26] and has been
117 shown to confer a RAS-independent constitutively activated state in the kinase domain [27]. This and
118 BRAF [L505H] have been reported in a small number of tumor entities including melanoma, prostate
119 and pancreatic cancer in cbioportal. BRAF [L505H] was shown to be activating mutation conferring
120 resistance to vemurafenib [28], and is characterized as oncogenic in OncoKB [29]
121 (https://www.oncokb.org/gene/, accessed 2023/12/13). The BRAF [L577I] mutation has been
122 reported in cutaneous melanoma and is characterized as variant with 'Unknown Oncogenic Effect' in
123 OncoKB.

124 In order to prioritize somatic mutations according to their putative oncogenicity, we next applied our
125 SVRACAS variant scoring pipeline (https://github.com/Jasonmbg/Simple.-Variant-Ranking-
126 Annotation-CAncer-Score, accessed 2022/10/24) to the 25,913 variants, leaving 4307 prioritized
127 variants (score ≥0.5) in 2969 genes (Supplementary Table 5). Amongst these, 54 matched with 72
128 CRC driver genes annotated IntOgen [30] (Fig. 1). The top frequently mutated genes in the ACCC tumors
129 were *APC* and *TP53* (Fig 1) and the presence of mostly loss-of-function mutations was in line with
130 previous observations [11, 31]. Next, we revealed differences in the 54 mutated genes that had matched
131 the IntOgen set, between the MSI and MSS CRC tumors of the ACCC cohort regarding somatic SNVs,
132 InDels and copy number alterations (CNAs). On the one hand, MSS tumors exhibited a higher
133 percentage of CNAs: Deletions of *SMAD2*, *SMAD4* and *MAP2K4*, and gains in *RNF6*, *BRCA2, NBEA* (chr
134 13q12-13) as well as in *GNAS* (chr 20q13.32) were only observed in MSS tumors. On the other hand,
135 frameshift and missense variants were more frequent (*RNF43*, *ACVR2A*, *TGFBR2*) or exclusive (*BCL9L*,
136 *PTPRC*) in the MSI cluster.

137 Mutational analysis implementing the COSMIC SBS signatures [32] revealed distinct patterns between
138 the MSI and MSS tumors (Supplementary Fig. 1). The majority of the MSI tumors were characterized
139 by SBS6, SBS26, and SBS20 signatures, all related to defective mismatch repair and, thus,
140 microsatellite instability. One sample (ACCC_CRC_15) was further characterized by signatures 10a
141 and 10b, which are related with mutations in *POLE*. Mutations in *POLE* are associated with a very
142 high TMB, which was indeed observed in this patient (375.994 mutations/Mb). On the contrary, MSS
143 samples were mainly enriched in the SBS1 and SBS5 signatures, which are related to age, stem-cell
144 division and "clock-like" processes, and are in line with a general low TMB in MSS tumors. However,
145 the DNA mismatch repair-associated signatures SBS6 and SBS15 were evident also in two MSS cases
146 (ACCC_CRC_05 & ACCC_CRC_06, respectively), possibly related to their high TMB. Patients
147 ACCC_CRC_07 and ACCC_CRC_18 were enriched with the unknown etiology signatures SBS93 and
148 SBS94, which have been identified in gastric and colorectal samples, respectively. Taken together,

149 these findings are in line with previous studies recapitulating the mutational landscapes of CRC,
150 where microsatellite instability is associated with a low frequency of CNAs and a high TMB [33, 34].

151 **Mapping mutational landscapes to functional networks**

152 To extend the analysis of mutational profiles in MSS and MSI CRC tumors, we utilized the public
153 datasets of the TCGA PanCan COAD/READ [11] and the CPTAC [24] cohorts. Complete datasets were
154 available for 434 patients (392 MSS, 42 MSI) from TCGA (Supplementary Table 2), and 79 patients (65
155 MSS, 14 MSI) from CPTAC (Supplementary Table 3). Highly scoring variants from SVRACAS annotation
156 (score ≥0.5) were identified in 7287 and 2588 genes in the TCGA and CPTAC datasets, respectively
157 (Supplementary Tables 6 and 7). Of those, only three genes had higher mutation frequencies in MSS
158 compared to MSI tumors: *APC* and *KRAS* genes were found more frequently mutated in MSS in both
159 public datasets, whereas mutations in *TP53* were significantly enriched only in the TCGA PanCancer
160 dataset. In contrast, 564 (TCGA) and 80 (CPTAC) genes had significantly higher mutation frequencies
161 in MSI compared to MSS tumors, and 42 genes were shared in both datasets. To uncover molecular
162 networks possibly related to microsatellite instability, we then performed functional enrichment
163 analysis by mapping the genes differentially mutated (MSS vs. MSI) in TCGA and CPTAC to Reactome
164 [35]. Eleven pathways were significantly enriched in MSI vs. MSS tumors in both, TCGA and CPTAC.
165 These signaling pathways were related to chromatin modification and organization, growth factor,
166 second messenger as well as Wnt signaling, and to RUNX1 transcriptional regulation (Fig. 2). In line
167 with only three genes having higher mutational frequencies in MSS vs. MSI, no Reactome pathway
168 was enriched in MSS tumors.

169 Next, we mapped the 2969 mutated genes from the ACCC CRC dataset to the 11 enriched Reactome
170 pathways. This identified 60 genes that were mutated with a higher frequency in ACCC MSI tumors
171 and in at least one of the two other CRC datasets. Functional and physical associations were
172 investigated by mapping these genes to the STRING database [36], identifying three protein interaction
173 networks with 60 genes and 154 edges (Supplementary Fig. 2). The first network comprised of genes
174 in angiogenesis as well as ERBB2, PI3K/AKT and PDGFR signaling. Network 2 was enriched in histone

175 modifications, NOTCH1, TCF and Wnt/β-Catenin signaling, while network 3 was characterized by
176 RUNX1 signaling, regulation of nucleotide-excision repair, double-strand break repair, and chromatin
177 remodeling. The Reactome and STRING network topologies mirrored prior knowledge of MSI CRC
178 thus further validating the ACCC CRC cohort and the classification of patients into MSS and MSI
179 subtypes [11, 12].

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

181 Aiming to estimate the activity of cancer-relevant signaling pathways in MSS and MSI groups of the
182 three CRC datasets (TCGA PanCan COAD/READ [11], CPTAC [24] and ACCC CRC), we next estimated the
183 activity of 14 signaling pathways based on transcriptomic data using the PROGENy computational
184 pipeline [37]. The activity of only the TGFβ signaling pathway was significantly higher in MSS than MSI
185 tumors in all three cohorts (Fig. 3A). This is in line with a recent report suggesting that TGFβ signaling
186 in MSS tumors stimulates cancer associated fibroblasts and establishment of an immune-excluded
187 tumor microenvironment [38]. In contrast, six pathways were stronger activated in MSI tumors in all
188 three cohorts (Fig. 3A). JAK-STAT signaling was the most activated pathway in MSI and has been
189 implicated with positive immunotherapy outcomes [39]. Then, kinase and pathway activity analysis was
190 performed using the phospho-proteomics measurements from the CPTAC cohort. This suggested the
191 cAMP-dependent protein kinase and TGFβ signaling pathways as more active in MSS tumors
192 compared to MSI (Supplementary Fig. 3A), confirming the PROGENy results for TGFβ. Activation of
193 Protein kinase A might be functionally related to gains and correlated expression of *GNAS* ($r=0.45$, p-
194 value < 0.05) found specifically in MSS tumors also in the CPTAC cohort (Supplementary Fig. 3C, D).
195 This pathway has been proposed as targetable driver mechanism in lung cancer [40]. In contrast,
196 several MAP kinases and PI3K signaling were significantly more activated in the MSI tumors
197 (Supplementary Fig. 3A, B). The latter findings confirm that MSI tumors exhibit a higher activation of
198 stress-response, inflammation/immune-modulation, and proliferation.

199 Next, we performed transcription factor (TF) activity analysis in MSS vs. MSI tumors and found 14 TFs
200 that had consistently different activity profiles in all three datasets. Six transcription factors including

201 RUNX2 (Runt-related transcription factor) were stronger activated in MSS CRCs compared to MSI in
202 all datasets (Fig. 3B). RUNX transcription factors are downstream effectors of TGFβ, thus matching
203 the results from pathway analysis. These TFs regulate basic cellular and developmental processes,
204 stem cell biology, and tumorigenesis [41]. In colorectal cancer, RUNX2 influences MAP kinase signaling
205 via regulation of multiple RTKs, and its absence, can lead to resistance against MEK inhibitors along
206 with its cofactor CBFB [42]. MSI CRCs presented significantly higher activities for the MYC, STAT2,
207 HIF1A, TP53, SPI1, IRF1 and IRF2 transcription factors (Fig. 3B). Some of these have previously been
208 associated with CRC and potentially impact immunogenicity, proliferation, DNA repair mechanisms
209 and cellular stress responses [43]. Collectively, pathway and transcription factor activation-profiling
210 point at elevated inflammatory processes in MSI tumors while tumors of the MSS type appear to be
211 associated with stem-cell features as well as an immune-excluded microenvironment.

212 **Transcriptomic stratification and immune deconvolution of ACCC-CRC samples**

213 We then estimated altered pathway and TF activities at the individual patient level to achieve a
214 personalized view on the ACCC patients. To this end, we used the scaled gene-expression values of
215 every patient as input and focused on the seven molecular pathways and 14 TFs, which we had found
216 significantly altered in the three cohort-based comparison of MSS vs. MSI CRC tumors. The resulting
217 estimated pathway and TF activities grouped the ACCC CRC patients into three patient clusters (Fig.
218 4A). To gain further insights into the underlying mechanisms that characterize the identified clusters,
219 we performed single sample Gene Set Enrichment Analysis (ssGSEA) [44], immune deconvolution [45],
220 and inference of Consensus Molecular Subtypes (CMS) [46].

221 The first patient cluster (Cluster_ID 1, n=13 patients) mostly included MSS (n=12) and CMS2 and
222 CMS3 cancers (Fig. 4A). In line with MSS tumors in the TCGA and CPTAC (Fig. 3A), the tumors in
223 cluster 1 were characterized by a low immunogenic behavior with sparse activation of the JAK-STAT
224 and NFkB inflammatory pathways, no enrichment of immune related hallmarks, and flat immune
225 infiltration scores (Fig. 4B). These findings are in line with the poor response MSS patients mostly
226 have to immunotherapies. Patient ACCC_CRC_15, the only MSI sample in this cluster, also showed an

227 overall trend of under-activation in the majority of estimated pathway and TF activities, whereas the
228 EGFR signaling cascade and MYC TF were found significantly over-activated in this *POLE*-mutant
229 tumor. This tumor had been diagnosed as serrated mucinous adenocarcinoma, which might explain
230 the separate clustering from the other MSI tumors [47]. The second cluster (Cluster_ID 2, n=7)
231 contained only MSS and mostly CMS4 samples (n=6) which were characterized by an overall
232 significant activation of the TGFβ signaling pathway, expression of genes associated with epithelial to
233 mesenchymal transition (EMT) and angiogenesis as well as activation for the SPI1, RUNX2 and
234 PRDM14. Almost all samples in the second cluster exhibited low activity of E2F4 and MYC,
235 accompanied with reduced cell cycle activity and unfolded protein response (Fig. 4A, B). This suggests
236 that tumors in the second cluster are characterized by an overall anti-proliferative capacity and a
237 more invasive phenotype, which could be related to the molecular and clinical characteristics of the
238 mesenchymal subtype (CMS4) [48]. The third cluster (Cluster_ID 3, n=8) comprised four MSI and four
239 MSS samples. It was mainly composed of CMS1 and CMS2 tumors with the exception of sample
240 ACCC_CRC_21 that was predicted as CMS3. Tumors in cluster 3 were characterized by a relatively
241 stronger activation of inflammatory (JAK-STAT, NFkB) and mitogenic (EGFR and MAPK) signaling
242 pathways, compared to patients in the other clusters. Moreover, MYC and E2F4 seemed to be
243 strongly activated. Conversely, a group of TFs, mainly associated with chromatin remodeling
244 (PRDM14, ASCL1, RUNX2) and developmental processes (HNF4A, TCF4) displayed a reduced activity
245 level compared to the other two clusters. This was most evident in the MSI samples of the third
246 cluster, while MSS samples in this cluster showed a more variable activation pattern. Collectively,
247 patients forming the third cluster reflected a more immune-reactive and less aggressive molecular
248 profile.

249 The MSS tumors showed an overall higher degree of heterogeneity, which was associated with the
250 prevalence of *KRAS* and *BRAF* mutations. *KRAS*-mutated MSS samples showed a trend towards
251 having lower activation of immune response pathways, i.e. having significantly lower activities in JAK-
252 STAT and NFkB pathways relative to the RAS_RAF_GNAS_wt MSS samples in the second patient

253 cluster (Fig. 4A). This was corroborated by ssGSEA findings and immune deconvolution estimates (Fig.
254 4B), suggesting that *KRAS*-mutant MSS tumors in the first patient cluster were less immune reactive
255 compared to the MSS tumors in the other clusters. These findings are in line with the observed low
256 activity of IRF1 and IRF2, regulators of the interferon response [49], in the majority of the *KRAS*-mutant
257 MSS samples of the first cluster (Fig. 4A). This is consistent with previous studies, which have
258 identified a mutant *KRAS-IRF2* molecular cascade, promoting an immunosuppressive environment in
259 CRC and fostering immunotherapy resistance [50]. Conversely, the transcriptomic circuits of MSS CRC
260 tumors in the third patient cluster showed greater similarity to the respective circuits in the MSI
261 tumors within that cluster, compared to the MSS samples in the other clusters. In summary, while
262 the evaluated transcriptomic programs primarily differentiated MSS and MSI samples and
263 corresponded with the prevalence of identified CMS subtypes, they also exposed differentially
264 enriched patient-cluster and patient-specific gene expression networks.

265 **From Oncogenic Mutational Landscapes and Transcriptomic Features Towards Personalized**
266 **Therapeutic Strategies**

267 Having seen higher functional heterogeneity in MSS compared to MSI tumors and that only few
268 pathways seemed to be consistently stronger activated in MSS tumors, we hypothesized that
269 precision oncology approaches might be particularly useful for MSS CRC. Along these lines, we went
270 out to unravel actionable alterations for individual ACCC CRC patients and interrelate them with their
271 respective transcriptomic profiles. This should provide more holistic views on potential vulnerabilities
272 and frame them as comprehensive personalized cancer maps, placing our special emphasis on MSS
273 cases. For this purpose, the SVRACAS prioritized somatic variants in each patient were annotated
274 with clinical actionability-information from the OncoKB precision oncology database [29]. This analysis
275 prioritized 34 actionable variants in 20 genes, reflecting these variants as predictive biomarkers of
276 drug sensitivity or resistance to particular targeted therapies (Fig. 5A).

277 Next, we assigned the ACCC CRC patients to seven MASTER interventional baskets [51], based on the
278 mutational profiles the respective tumors presented with. The MSI-status and TMB were regarded

279 for assignments to the immune evasion (IE) basket. Seventeen patients of the ACCC cohort (~63%)
280 were stratified into at least one interventional basket (Fig. 5B). Tumors in patient cluster 1 were
281 mainly characterized by overrepresentation of the RAS-MEK-ERK (RME, 8 out of 9 patients), the PI3K-
282 AKT-mTOR (PAM, 3 out of 9 patients), and the immune evasion (IE, 3 out of 9 patients) baskets. An
283 interventional basket (i.e., RME) could be assigned to just one (ACCC_CRC_16) out of seven MSS
284 patients forming cluster 2, in line with the rather low TMB in the tumors forming this cluster. Cluster
285 3 was mostly represented by RME (5 out of 7 patients) and PAM as well as IE (4 out of 7 patients
286 each) baskets.
287 The five MSI- and two MSS-patients could qualify for treatment with immune checkpoint-inhibitors
288 (e.g., pembrolizumab), the MSS-patients in the light of the high TMB in the tumors (Fig. 5B). Two MSI
289 and another two MSS tumors carried BRAF V600E mutations and might benefit from treatment with
290 encorafenib and cetuximab [52]. Seven patients carried *KRAS* or *NRAS* mutations, rendering those
291 patients incompatible with EGFR-antibody therapy [53]. Inhibitors specific for the most common KRAS
292 variant G12D, which was present in three cases, are currently in preclinical and clinical development
293 (e.g. MRTX1133, NCT05737706). Tumors carrying druggable KRAS G12C mutations [3] were not present
294 in our cohort. No specific treatments were clinically available for the patients carrying other RAS
295 mutations, i.e., ACCC_CRC_13 (G12V) and ACCC_CRC_27 (G12S). However, also non-specific Pan-
296 KRAS inhibitors are in development [54], which might be applied once having proven efficacious.
297 Hence, molecular profiling identified potential treatment options for the majority of patients in the
298 cohort and more than one interventional basket was mapped to eight patients, suggesting a
299 combination of targetable driver mechanisms in the respective patients.

300 We finally went ahead with two patients of the MSS type to prove the utility of our approach
301 towards suggesting potential vulnerabilities for tailored therapeutic interventions. Patients
302 ACCC_CRC_05 and ACCC_CRC_06 had presented with adenocarcinomas of the CMS3 subtype and
303 had a right-sided location. Patient ACCC_CRC_05 (Fig. 5C) was characterized by an ERBB2 p.R678Q
304 mutation, which has a modest biological rationale for treatment with trastuzumab-deruxtecan [55]

305 (NCT evidence level m4). Furthermore, this patient carried an ATM p.F61Lfs*15 mutation, which has
306 been found in few, mostly colorectal tumors (CosmicID: COSV53735143, accessed 2023/12/14).
307 While preclinical evidence supports PARP inhibition [56] (NCT evidence level m3) this was recently
308 shown (NCT02693535) to be ineffective in *ATM*-mutated CRC [57] thus not giving it high priority. In
309 patient ACCC_CRC_05, we identified *PTEN* p.T319* and TSC2 p.R1451Afs.25 mutations, both of which
310 have biological justification and preclinical evidence [58] suggesting treatment with mTOR inhibitors
311 (NCT evidence level m3-m4). Tumor gene expression analysis for this patient demonstrated an over-
312 activation of PI3K/AKT/mTOR signaling. The data further suggested an under-activation in the EGFR
313 and JAK-STAT signaling pathways, which might indicate a predominantly PI3K/mTOR-driven cancer
314 phenotype for this patient, further supporting the rationale for mTOR inhibition.

315 MSS patient ACCC_CRC_06 carried several potentially druggable alterations (Fig. 5C). The BRAF
316 V600E mutation in this tumor could be targeted with Encorafenib and Cetuximab as an approved
317 combination for second line targeted therapy after chemotherapy [1]. BRCA1 p.K654Sfs*47 and BRCA2
318 p.T3085Qfs*19 mutations in this patient suggested PARP inhibition as a potential option. Tumor
319 transcriptomic analysis revealed an over-activation of the PI3K/AKT/mTOR and under-activation of
320 the DNA repair and JAK/STAT molecular circuits. A frameshift mutation in *TSC2* (p.N1017Tfs*13)
321 would suggest mTOR inhibition as therapeutic strategy [59]. However, in contrast to patient
322 ACCC_CRC_05, we noticed a trend towards activation of the MAPK signaling pathway in patient
323 ACCC_CRC_06, which might be associated with the presence of the BRAF V600E mutation. Thus, the
324 value of targeting mTOR in the context of an activating BRAF mutation seems questionable [60].

325 Patients ACCC_CRC_05 & _06 both presented with high TMB, which has been observed in MSS
326 tumors before [31], and mutational signatures associated with DNA mismatch repair (SBS6 and SBS15
327 in patients ACCC_CRC_05 and ACCC_CRC_06, respectively). While the TMB might indeed indicate
328 application of immunotherapies, the lack of mutations in *POLE* [61], and the modest stromal and
329 immune signatures in these tumors were not supportive of this (Fig. 4), suggesting immune
330 checkpoint-inhibition as an only hypothetical treatment option [62]. Yet, our analysis pointed at

331 rationale treatment options for these two as well as for most other patients in the ACCC CRC cohort.
332 Collectively, our study underlines the significant value of incorporating various molecular and
333 functional layers to support therapeutic decision-making.

334 Discussion

335 Here, we set out to characterize the molecular and activity landscapes MSS and MSI colorectal
336 cancer. By implementing an integrative computational workflow we analyzed tumors from 28
337 patients of a prospective CRC patient cohort of the Athens Comprehensive Cancer Center (ACCC), in
338 combination with two large public multi-omics datasets. There, we explored the molecular profiles of
339 MSS vs. MSI CRCs, their relation to *KRAS*/*BRAF* mutations, and pathway as well as transcription factor
340 activities in the context of personalized medicine. Firstly, we observed novel associations between
341 particular genes and their increased mutational frequencies specifically in MSI tumors. Mutations
342 were enriched in genes that are involved in molecular pathways mainly related to transcriptional
343 regulation cascades by RUNX1, chromatin modification mechanisms, and in signaling cascades that
344 are related to Reactome 'fusion mutants/proteins' (i.e., FGFR1/FLT3). These findings resonate with
345 research indicating the central role of RUNX1 in the malignant transformation process and metastasis
346 of CRC through mechanisms, like the Wnt/β-catenin signaling pathway and EMT [63, 64]. Fusion events
347 involving FGFR have been considered as therapeutic targets in various solid tumors [65]. In contrast,
348 MSS tumors demonstrated a pronounced enrichment in CNAs. Deletions were recurrently found in
349 *SMAD2*, *SMAD4*, and *MAP2K4*, which are pivotal in the TGFβ and JNK signaling pathways [66, 67],
350 respectively. Gains were observed in chromosomes 13q12-13 and 20q13.32, which have been
351 associated with early events in the development of colorectal cancer [68], also affecting *GNAS*. These
352 CNAs exclusively presented in MSS, suggesting a unique MSS genomic signature with implications in
353 MSS disease etiology and therapeutic targeting [69].

354 Secondly, our comprehensive multi-omics analysis indicated differentially regulated circuits at the
355 transcriptional level. Immune related biological mechanisms, such as JAK-STAT and NF_KB signaling

356 pathways, were found upregulated in MSI tumors, consistent with the higher 'immunogenicity' in the
357 context of microsatellite instability [70]. Conversely, the TGFβ pathway showed greater activation in
358 the MSS tumors and has been associated with epithelial-to-mesenchymal transition (EMT),
359 metastasis, and therapy resistance [71, 72]. We further observed that the synchronous presence of *KRAS*
360 mutations with microsatellite instability had a stronger effect on the immune profile of CRC tumors
361 than mutations in *BRAF*. Hence, we identified context-specific functional associations that are likely
362 related to the higher functional heterogeneity within MSS CRC tumors. Previous studies have
363 emphasized the intricate role *KRAS* mutations play in shaping the tumor microenvironment (TME) of
364 colorectal cancer. These mutations seem to foster an immunosuppressive environment, which is
365 marked by a diminished anti-tumor immune response [73].

366 Next, we identified transcription factors with differential activity between MSS and MSI CRC tumors.
367 MSI CRC tumors showed higher activities of MYC, STAT2, HIF1A, TP53, SPI1, IRF1, and IRF2. These
368 factors are instrumental in various cellular processes, such as cell growth, response to stress, oxygen
369 homeostasis, and immune modulation [43]. Hence, therapeutic regulation of TFs might be a promising
370 strategy for CRC therapy owing to their underlying role as key effectors of various signal transduction
371 and tumor-associated immune responses [74]. Distinct transcriptional profiles were observed in MSS
372 CRCs, indicating the heightened activity of RUNX2 upstream of TGFβ signaling [75]. Yan et al. revealed
373 that RUNX2 is critical for the maintenance of the stem cell-like properties of CRC cells as well as for
374 the promotion of CD44-induced EMT in CRC[76]. Collectively, from our integrative analysis we noticed
375 the significant deregulation of specific members of the RUNX TF family at DNA (*RUNX1*) or mRNA
376 (*RUNX2*) levels, supporting their emerging role as prognostic biomarkers in CRC [77].

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

382 clonal evolution and tumorigenesis of CRC [78]. Indeed, this aligns with a recent study proposing a
383 novel molecular classification of CRC tumors based on single-cell sequencing data [79]. This study
384 identified a subgroup of MSS samples (iCMS3_MSS), which exhibited more similarity with the
385 transcriptional and genomic levels of MSI CRC tumors (iCMS3_MSI), forming a novel epithelial
386 subtype [79]. This may indicate that the MSI status, even though clearly distinct at the genomic level,
387 may be viewed as a continuum through different states and sub-conditions at the transcriptomic
388 layer [80]. The potential therapeutic consequences of this, however, remain to be uncovered.

389 Finally, we sought to provide a personalization of molecular findings, as the identification of
390 targetable alterations is a primary focus of current clinical research, to facilitate personalized
391 treatment strategies. Effective therapeutic options exist for MSI CRCs employing immune checkpoint
392 inhibitors. In contrast, the management of MSS cancers still predominantly relies upon conventional
393 chemotherapies, as the most frequently encountered mutations in MSS CRC affect non-druggable
394 tumor suppressor genes, such as *APC* and *TP53*. Indeed, the prevalence of well-established
395 targetable mutations with clinical evidence supporting therapeutic interventions is exceedingly low in
396 CRC. These primarily encompass BRAF V600E mutations, which we observed also in our cohort and
397 have proven more difficult to target in CRC compared, for example, to melanoma [52]. Clinical
398 treatment options are often limited even when mutations in potentially druggable drivers have been
399 identified [81]. Hence, the integration of expression data might prove beneficial in clarifying
400 therapeutic decisions for patients with uncertain mutation-driven treatment options, which we
401 illustrated with patients ACCC_CRC_05 and ACCC_CRC_06. Our integrated analysis suggested
402 treatment options for these patients (PARP/mTOR inhibition) that were unknown when the patients
403 were indeed in clinical treatment.

404 The added value of integrating molecular modalities to foster precision medicine was initially
405 introduced with the WINTHER clinical study [82] (NCT01856296). WINTHER was the first trial to
406 describe the combination of matched genomic and transcriptomic information in metastatic tumors,
407 showcased the significance of integrated molecular profiling into improved therapeutic

408 recommendations, and expanded the portfolio of putative therapeutic interventions [82]. Likewise, the
409 more recent MASTER/NCT clinical study has illustrated that the integration of WES/WGS with RNA
410 sequencing can result in a clinical benefit and boost patient treatment stratification in patients with
411 advanced cancer of rare cancer entities [51]. Our findings support these concepts, while we placed our
412 focus on MSS and MSI CRC, in the context of the mutation status in the *KRAS* and *BRAF* genes. Our
413 study initially suffered from the relatively small sample size of the retrospective ACCC cohort, which
414 impeded direct generalization of inferential statistics and related biological comparisons. To cope
415 with this limitation, we took advantage of public data sets of the TCGA and CPTAC cohorts,
416 constructed an integrative workflow, and comprehensively profiled the sequencing data from the
417 ACCC and the public multi-omics cancer datasets. Collectively, this allowed us to perform a patient-
418 centric analysis capitalizing on the available molecular information as a guide to therapeutic
419 interventions. In conclusion, our approach integrates mutation ranking with immune deconvolution
420 and pathway as well as transcription factor activity analysis, to suggest personalized treatment
421 strategies. In the future, this will likely be complemented with proteomic information [83] and holds
422 the promise for enhancing the biological rationale behind individualized therapeutic interventions
423 even it does not constitute a therapeutic protocol per se. Yet, it serves as a paradigm that supports
424 and augments the concept of tailored medicine in oncology.

425 Methods

426 Patients & clinical specimens

427 Colorectal cancer patients operated in the surgical department of Gennimatas General Hospital in
428 Athens, which is part of the Athens Comprehensive Cancer Center (ACCC), participated in the study.
429 The study was approved by the ethics committee of the Gennimatas Hospital (protocol number
430 27659). Written informed consent was obtained from all patients. Only patients (n=28) with
431 histologically confirmed primary colorectal adenocarcinoma and one case of rectal cancer who had
432 received radiotherapy before surgery, were enrolled in the study cohort. Specimens from each

433 patient were obtained from both the primary tumor and corresponding normal mucosa, which was
434 taken from the histologically negative margins (at least 5 cm away from the macroscopically defined
435 cancerous area) to ensure the absence of cancer cells. Upon surgery, resected samples were placed
436 in 5-mL RNA*later* (Ambion, Austin, TX, USA) to preserve RNA integrity.

437 **DNA and RNA extraction**

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

444 **Whole Exome Sequencing (WES) and variant calling**

445 DNA libraries were prepared from tumor and matched normal tissues of ACCC patients using Low
446 Input Exome-Seq Human v6 (Agilent Technologies, Waldbronn, Germany) for exome capture and
447 analyzed on a NovaSeq 6000 (Paired-End 150bp S2, Illumina, San Diego, CA, USA) in the DKFZ High-
448 Throughput Sequencing Core Facility. The resulting raw sequencing reads (median: 123 million
449 reads/sample; spread: 80-148 million reads/sample) were preprocessed based on the internal WES
450 alignment workflow (AlignmentAndQCWorkflows: 1.2.73-1), as part of the DKFZ-ODCF OTP pipeline
451 [84]. Briefly, after initial quality control using FastQC, raw reads were aligned to a hg19 customized
452 version of the 1000 genomes reference genome
453 (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/phase2_reference_assembly_sequen
454 ce/hs37d5.fa.gz) (1KGRef_PhiX), using bwa-mem (version 0.7.15). PhiX-genome sequence was added
455 to the reference as an additional contig to remove spike-in sequences used to identify potential
456 cross-contaminations and sample swaps. Marking of duplicated reads was performed with

457 Sambamba tools (https://github.com/biod/sambamba/releases, version 0.6.5) embedded in the
458 DKFZ-ODCF Roddy pipeline (https://github.com/DKFZ-ODCF/AlignmentAndQCWorkflows).

459 High-confidence functional somatic single nucleotide variants (SNVs) and InDels (Insertions &
460 Deletions) were identified using DKFZ-ODCF SNVCallingWorkflow (https://github.com/DKFZ-
461 ODCF/SNVCallingWorkflow, version 1.2.166-1) and IndelCallingWorkflow (https://github.com/DKFZ-
462 ODCF/IndelCallingWorkflow, version 1.2.177-0), respectively, based on the BAM files of each tumor
463 and its paired control sample. Various quality control criteria were performed to exclude spurious
464 or/and low quality calls (such as identification of low quality bases, variant allele fraction and
465 presence of strand bias artifacts), finally assigning a confidence value to each variant. Sample
466 ACCC_CRC_02 was disregarded in further analysis as COSMIC SBS signature analysis (v3.4-October
467 2023) flagged this as having a potential PCR/sequencing artifact-based strand bias (prevalence of C>A
468 alterations indicated presence of COSMIC SBS52; https://cancer.sanger.ac.uk/signatures/sbs/sbs52/,
469 accessed 2023/12/21). Finally, the ANNOVAR tool [85] (version February 2016) for variant annotation
470 was utilized for the derivation of all functional mutations defined only by the exonic mutations
471 including missense, nonsense, splice-site, and stop-loss alterations, by setting a confidence value of
472 at least 8 out of 10. High-confidence InDels were selected using the Platypus tool [86] implemented in a
473 custom workflow (https://github.com/DKFZ-ODCF/IndelCallingWorkflow, version 0.8.1.1).

474 **SVRACAS variant annotation and gene prioritization**

475 To efficiently prioritize the resulting high-confidence somatic variants for each patient with the
476 ultimate goal to identify informative oncogenic alterations, the SVRACAS custom variant scoring
477 pipeline was applied (https://doi.org/10.5281/zenodo.5636747, accessed 2022/10/24). This
478 integrative workflow consists of two major steps: variant annotation and prioritization of somatic
479 functional alterations. Briefly, the SVRACAS ranking module integrates 15 biological resources using
480 the open-source OpenCRAVAT/OakVar comprehensive workflow for cancer variant annotation and
481 exploitation [87]. The aggregated output has four evidence categories: "Pathogenicity evidence"
482 (spanning variant effect prediction tools), a "Cancer evidence" layer (covering cancer variant

483 interpretation knowledge bases), a "Clinical support" indication (nominating clinically relevant
484 alterations), and "Expression information" indicating whether an affected gene is expressed (i.e., ≥ 10
485 read counts in at least 5 samples). An integrated scoring value was then assigned to each variant as a
486 total SVRACAS score (https://github.com/Jasonmbg/Simple.-Variant-Ranking-Annotation-CAncer-
487 Score). This holistic prioritization scheme ranges from 0 to 1, representing an average approximation
488 of the four different types of evidence, based on single nucleotide changes and InDels that occur in
489 the protein-coding space. Here, a cutoff value ≥ 0.5 was used to include only variants annotated in at
490 least two of the four evidence categories. Overall, the selection of individual relative cutoffs and
491 weighted importance of different sources of evidence was based on proposed standard operating
492 procedures that built on guidelines from cancer consortia such as VICC [88] and ComPerMed [89]. Protein-
493 coding genes were then mapped to the MANE Select Set (version 0.9) of matched representative
494 transcripts [90]. Genes that did not map in MANE Select were next assessed for predicted
495 consequences respective mutations have according to the Sequence Ontology [91]. The aberrant
496 transcript having the highest predicted impact was selected. The longest transcript was chosen when
497 the same impact was annotated by Sequence Ontology for several transcripts representing the same
498 gene. Genes in the ACCC CRC dataset with SVRACAS prioritized mutations were mapped to a curated
499 list of CRC cancer driver genes in the Integrative OncoGenomics database (IntOGen) [30]
500 (https://www.intogen.org, v1-2020).

501 **Identification and annotation of Copy Number Alterations**

502 Copy Number Alterations (CNAs) were inferred from the whole exome sequencing data using a
503 customized workflow (Supplementary File 1). Briefly, CNVkit (version 0.9.3) [92] was utilized with
504 default parameters to infer discrete copy number segments. Next, tumor cell content and ploidy
505 were estimated using a method adapted from ACEseq [93]. Absolute copy numbers and allele-specific
506 copy numbers as well as the decrease in heterozygosity (DH) [94] were estimated segment-wise, for
507 each possible combination of Tumor cell content and ploidy.

508 **RNA-Sequencing and analysis**

509 RNAseq libraries were prepared using the TruSeq Stranded library protocol (Illumina) and sequencing
510 was done with a NovaSeq 6000 (Paired-end 100 bp S1) in the DKFZ Sequencing Core Facility to collect
511 a median of 75 million reads/sample (spread: 65-99 million). Processing of RNA-Seq data was
512 performed using a customized RNAseq Workflow as a part of the DKFZ – ODCF Roddy pipeline
513 (https://github.com/DKFZ-ODCF/RNAseqWorkflow, version 1.3.0). After initial QC, the RNAseq raw
514 reads were aligned to a customized hg19 version of the 1000 genomes reference genome
515 (1KGRef_PhiX):
516 (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/phase2_reference_assembly_sequen
517 ce/hs37d5.fa.gz), by a 2 pass alignment using STAR aligner [95] (version 2.5.3a) and gencode 19 gene
518 models. PhiX-genome sequence was added to the reference as an additional contig to remove spike-
519 in sequences that had been included to find potential cross-contaminations and sample swaps.
520 Additional QC after alignment was performed using samtools *flagstat* command and the RNA-SeQC
521 tool [96] (version 1.1.8). The featureCounts function [97] as part of the Subread software package [98]
522 (version 1.5.1) was used to determine the respective raw counts per gene, using gencode 19 gene
523 models. Downstream RNASeq analysis was performed with R (R version 4.1.0)/Bioconductor
524 software. The R package EnsDb.Hsapiens.v75 was utilized to annotate gene symbols. Non-specific
525 intensity filtering was performed to exclude genes with less than 10 read counts in less than five
526 samples. TMM normalization was conducted using the edgeR R package [99] (version 3.36.0), followed
527 by incorporating sample weights and increase statistical power using the Voom function. Finally, the
528 R package immunedeconv [45] (version 2.1.0) was utilized to estimate the composition and density of
529 immune and stromal infiltration, using the non-log-transformed Transcripts Per Kilobase Million
530 (TPM) values as input.

531 **TCGA and CPTAC multi-OMICs data**

532 Public data from two CRC cohorts were downloaded from the cBioPortal multi-omics cancer portal
533 (https://www.cbioportal.org/) i.e., the PanCancer Atlas TCGA Colorectal adenocarcinoma [11] and the
534 CPTAC Prospective Colon [100, 101] datasets. For both datasets, only samples were processed that had

535 been profiled for SNVs, CNAs, gene expression (RNAseq). Phosphoproteomic data were available only
536 from CPTAC. SNVs were retrieved and data analysis was performed using the R package maftools [102]
537 (version 2.10.0). The significance of differential mutational frequencies between the MSS and MSI
538 tumors within TCGA and CPTAC cohorts was assessed using Fisher's exact test based on a 2x2
539 contingency table (adjusted p-value ≤0.1, minimum number of samples harboring mutated genes
540 ≥4), exploiting the mafCompare function of maftools.

541 **Inference of microsatellite instability status in the CRC datasets**

542 We utilized the MSI MANTIS [103] estimators of microsatellite instability provided with the TCGA pan-
543 cancer dataset, using the following categorization: A TCGA tumor was classified as MSI if the
544 respective value from MSI MANTIS was higher than 0.6, while tumors with relative values lower than
545 0.4 were classified as MSS. Any tumors with relative values between 0.4 & 0.6 were denoted as
546 "Undetermined" and excluded from further analysis. Information about MSI-status of respective
547 samples was available also for the CPTAC dataset [101] and utilized as provided. Then, the PREMSIm R
548 package [23] (version 1.0) was employed to infer the microsatellite status based RNA-sequencing data
549 from all three datasets (TCGA, CPTAC, ACCC), using non-log-transformed Transcripts Per Kilobase
550 Million values from RNA-sequencing. MSI-L and MSS tumors were grouped together in all datasets
551 and the term MSS used for this group to distinguish from MSI tumors.

552 **Functional enrichment analysis**

553 Functional Enrichment analysis was performed using the R packages clusterProfiler [104] (version 4.2.2),
554 and ReactomePA [105] (version 1.38.0) was used for overrepresentation analysis based on REACTOME
555 molecular pathways. ssGSEA [44] (https://github.com/broadinstitute/ssGSEA2.0) was implemented
556 based on the ssGSEA projection methodology [106]. The hallmark signatures file (version 7.5.1) was
557 used in gmt format from the respective Molecular Signatures Database (MSigDB) [107] gene sets
558 (https://www.gsea-msigdb.org/gsea/msigdb/). The script ssgsea-cli.R

559 (https://github.com/broadinstitute/ssGSEA2.0) was run in a Windows 10 x64 (build 19044)
560 environment using command line and the following arguments:

561 *[-i ACCC.CRC.RNASeq.VoomCPM.TMM.Filtered -o ACCC -n none -d ./ssGSEA2.0-*
562 *master/db/msigdb/h.all.v7.5.1.symbols.gmt -w 0.75 -c z.score -p 1000 -m 10]*. Protein associations
563 were retrieved and visualized using STRING [36] (https://string-db.org, accessed 2023/12/04), and K-
564 means clustering was performed to identify functional protein networks.

565 **Pathway, TF and Kinase activity inference analyses**

566 The upper-quartile of processed and log2 transformed RSEM estimated counts from CPTAC RNAseq
567 data were normalized using the function *normalizeQuantiles* from the limma R package [108] (version
568 3.50.0). Then, a non-specific intensity filtering procedure was implemented to remove features that
569 were not expressed at least in the group with the lowest number of samples. The same
570 preprocessing approach was applied for the TCGA dataset, with two additional steps: as the
571 estimated raw RSEM counts had been batch effect-corrected, many features assigned by NA values
572 were removed. Then, the data were log2-transformed if the maximum value was >50. Differential
573 expression analysis in both datasets (MSI vs MSS) was applied with the limma R package [108].

574 The PROGENy (Pathway RespOnsive GENes R package [37] (version 1.16.0) was implemented to identify
575 putative differentially activated molecular pathways, based on the derived differential moderated
576 statistics extracted from the limma pipeline (i.e., top 100 most responsive genes for each pathway).
577 PROGENy infers the activity of 14 major signaling pathways, based on the exploitation of consensus
578 gene signatures and estimated from a large compendium of perturbation experiments [37]. The
579 DoRothEA-decoupleR pipeline [109] (version 1.6.0) was implemented to elucidate differentially
580 activated regulatory networks. Relative activities were inferred for every transcription factor (TF)
581 from the deregulated expression of its respective target genes, using the normalized weighted mean
582 statistical metric from the decoupleR R package [110] (version 2.1.6). Customized barplots were created

583 using the R package ggplot2 [111] (version 3.3.5) for the visualization of differentially altered TFs and
584 pathways.

585 Phosphoproteomics data from the CPTAC cohort were processed and analyzed to infer differentially
586 activated signaling cascades between MSS and MSI colon tumors. To this end, normalized
587 phosphosite ratios were used, representing scaled intensity values. Duplicated entries were removed
588 and any phosphosites were excluded that had NA values across more than 85% of the total samples.
589 Differential protein phosphorylation analysis was performed using the limma R pipeline (version
590 3.50.0). Omnipath-decoupleR [110] and Post-translational Modification Set Enrichment Analysis (PTM-
591 SEA) [112] with the PTMSigDB signaling pathways' collection (version 2.0.0) were implemented to
592 determine differentially activated kinases and signaling pathways, respectively. The Omnipath R
593 package [113] (version 3.2.8) was utilized to retrieve the prior-knowledge interactions composed by
594 kinase-target relationships, whereas the R package decoupleR [110] (version 2.1.6) was used to
595 estimate the relative biological activities by application of statistical methods (normalized weighted
596 mean).

597 **Patient clustering**

598 Unsupervised hierarchical clustering on principal components (HCPC) method was implemented,
599 based on the R packages FactoMineR [114] (version 2.8) and Factoshiny
600 (http://factominer.free.fr/graphs/factoshiny.html, version 2.4). The selected scaled pathway and TF
601 activities per patient were used as input for the *PCA* function. Then, hierarchical clustering on the
602 first 8 components (accounting ~95% of retained information) was performed using the *HCPC* (kk=Inf,
603 max=10, nb.clust = -1, consol = TRUE) function. Plots were created using the ComplexHeatmap R
604 package [115]. Patients were re-ordered within each cluster according to their microsatellite instability
605 status and selected mutational groups, namely BRAF_mut, KRAS_mut, NRAS_mut, GNAS_mut,
606 RAS_RAF_GNAS_mut, and RAS_RAF_GNAS_wt.

607 **Identification of actionable alterations for construction of personalized cancer patient maps**

608 Aiming to identify actionable vulnerabilities at the single patient resolution within the ACCC-CRC
609 cohort, the qualified somatic variants from the SVRACAS scoring scheme were integrated with clinical
610 implications from the OncoKB precision oncology database [29]. Variant oncogenic effect annotation
611 was applied to keep only variants characterized as oncogenic or likely oncogenic. Therapeutically
612 actionable variants in any cancer type were assigned with an OncoKB Therapeutic Level of Evidence
613 (https://www.oncokb.org/therapeutic-levels, accessed 2022/12/21). Genes carrying actionable
614 variants were assigned to the treatment baskets of the National Center for Tumor Diseases/German
615 Cancer Consortium (NCT/DKTK) MASTER (Molecularly Aided Stratification for Tumor Eradication
616 Research) program, that are based on a set of 472 genes which participate in biological processes
617 and cellular pathways with established therapeutic recommendations [51].

618 **Reporting summary**

619 Further information on experimental design is available in the Nature Research Reporting Summary
620 linked to this article.

621 Data availability

622 All sequencing data supporting the findings of this study have been uploaded to the European
623 Genome-Phenome archive (http://www.ebi.ac.uk/ega/), and are available with controlled access of a
624 DAC via the accession numbers EGAD00001011262 and EGAD00001011263. All computational code,
625 including R scripts, quarto documents and further information to reproduce the complete
626 bioinformatics analyses, are available on Zenodo
627 (https://zenodo.org/doi/10.5281/zenodo.10959699) and Github
628 (https://github.com/Jasonmbg/ACCC-CRC, https://github.com/Jasonmbg/TCGAPanCan-CRC and
629 https://github.com/Jasonmbg/CPTAC-Colon) repositories for open-access.

630 References

631 1. Kopetz S, *et al.* Encorafenib, Binimetinib, and Cetuximab in BRAF V600E-Mutated Colorectal
632 Cancer. *N Engl J Med* **381**, 1632-1643 (2019).

633
634 2. Casak SJ, *et al.* FDA Approval Summary: Tucatinib with Trastuzumab for Advanced
635 Unresectable or Metastatic, Chemotherapy Refractory, HER2-Positive RAS Wild-Type
636 Colorectal Cancer. *Clin Cancer Res* **29**, 4326-4330 (2023).

637
638 3. Janne PA, *et al.* Adagrasib in Non-Small-Cell Lung Cancer Harboring a KRAS(G12C) Mutation.
639 *N Engl J Med* **387**, 120-131 (2022).

640
641 4. Siegel RL, Wagle NS, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2023. *CA*
642 *Cancer J Clin* **73**, 233-254 (2023).

643
644 5. Sung H, *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality
645 Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209-249 (2021).

646
647 6. Scharpf RB, *et al.* Genomic Landscapes and Hallmarks of Mutant RAS in Human Cancers.
648 *Cancer Res* **82**, 4058-4078 (2022).

649
650 7. Dang HX, *et al.* The clonal evolution of metastatic colorectal cancer. *Sci Adv* **6**, eaay9691
651 (2020).

652
653 8. Househam J, *et al.* Phenotypic plasticity and genetic control in colorectal cancer evolution.
654 *Nature* **611**, 744-753 (2022).

655
656 9. Loree JM, *et al.* Classifying Colorectal Cancer by Tumor Location Rather than Sidedness
657 Highlights a Continuum in Mutation Profiles and Consensus Molecular Subtypes. *Clin Cancer*
658 *Res* **24**, 1062-1072 (2018).

659
660 10. Linares J, *et al.* Long-term platinum-based drug accumulation in cancer-associated fibroblasts
661 promotes colorectal cancer progression and resistance to therapy. *Nat Commun* **14**, 746
662 (2023).

663
664 11. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and
665 rectal cancer. *Nature* **487**, 330-337 (2012).

666
667 12. Mendelaar PAJ, *et al.* Whole genome sequencing of metastatic colorectal cancer reveals prior
668 treatment effects and specific metastasis features. *Nat Commun* **12**, 574 (2021).

669
670 13. de la Chapelle A, Hampel H. Clinical relevance of microsatellite instability in colorectal cancer.
671 *J Clin Oncol* **28**, 3380-3387 (2010).

672
673 14. Hause RJ, Pritchard CC, Shendure J, Salipante SJ. Classification and characterization of
674 microsatellite instability across 18 cancer types. *Nat Med* **22**, 1342-1350 (2016).

675
676 15. Boland CR, *et al.* A National Cancer Institute Workshop on Microsatellite Instability for cancer
677 detection and familial predisposition: development of international criteria for the

678 determination of microsatellite instability in colorectal cancer. *Cancer Res* **58**, 5248-5257
679 (1998).

680

681 16. Lochhead P*, et al.* Microsatellite instability and BRAF mutation testing in colorectal cancer
682 prognostication. *J Natl Cancer Inst* **105**, 1151-1156 (2013).

683

684 17. Poynter JN*, et al.* Molecular characterization of MSI-H colorectal cancer by MLHI promoter
685 methylation, immunohistochemistry, and mismatch repair germline mutation screening.
686 *Cancer Epidemiol Biomarkers Prev* **17**, 3208-3215 (2008).

687

688 18. Lynch HT, Snyder CL, Shaw TG, Heinen CD, Hitchins MP. Milestones of Lynch syndrome: 1895-
689 2015. *Nat Rev Cancer* **15**, 181-194 (2015).

690

691 19. Le DT*, et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* **372**,
692 2509-2520 (2015).

693

694 20. Innocenti F*, et al.* DNA Mutational Profiling in Patients With Colorectal Cancer Treated With
695 Standard of Care Reveals Differences in Outcome and Racial Distribution of Mutations. *J Clin*
696 *Oncol* **42**, JCO2300825 (2023).

697

698 21. De Smedt L*, et al.* Microsatellite instable vs stable colon carcinomas: analysis of tumour
699 heterogeneity, inflammation and angiogenesis. *Br J Cancer* **113**, 500-509 (2015).

700

701 22. Cleyle J*, et al.* Immunopeptidomic analyses of colorectal cancers with and without
702 microsatellite instability. *Mol Cell Proteomics* **21**, 100228 (2022).

703

704 23. Li L, Feng Q, Wang X. PreMSIm: An R package for predicting microsatellite instability from the
705 expression profiling of a gene panel in cancer. *Comput Struct Biotechnol J* **18**, 668-675 (2020).

706

707 24. Edwards NJ*, et al.* The CPTAC Data Portal: A Resource for Cancer Proteomics Research. *J*
708 *Proteome Res* **14**, 2707-2713 (2015).

709

710 25. Zampino G*, et al.* Diversity, parental germline origin, and phenotypic spectrum of de novo
711 HRAS missense changes in Costello syndrome. *Hum Mutat* **28**, 265-272 (2007).

712

713 26. Dankner M, Rose AAN, Rajkumar S, Siegel PM, Watson IR. Classifying BRAF alterations in
714 cancer: new rational therapeutic strategies for actionable mutations. *Oncogene* **37**, 3183-
715 3199 (2018).

716

717 27. Hu J*, et al.* Allosteric activation of functionally asymmetric RAF kinase dimers. *Cell* **154**, 1036-
718 1046 (2013).

719

720 28. Choi J*, et al.* Identification of PLX4032-resistance mechanisms and implications for novel RAF
721 inhibitors. *Pigment Cell Melanoma Res* **27**, 253-262 (2014).

722
723 29. Chakravarty D*, et al.* OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol* **1**, 1-16
724 (2017).

725
726 30. Martinez-Jimenez F*, et al.* A compendium of mutational cancer driver genes. *Nat Rev Cancer*
727 **20**, 555-572 (2020).

728
729 31. Jan YH, Tan KT, Chen SJ, Yip TTC, Lu CT, Lam AK. Comprehensive assessment of actionable
730 genomic alterations in primary colorectal carcinoma using targeted next-generation
731 sequencing. *Br J Cancer* **127**, 1304-1311 (2022).

732
733 32. Alexandrov LB*, et al.* The repertoire of mutational signatures in human cancer. *Nature* **578**,
734 94-101 (2020).

735
736 33. Dienstmann R, Vermeulen L, Guinney J, Kopetz S, Tejpar S, Tabernero J. Consensus molecular
737 subtypes and the evolution of precision medicine in colorectal cancer. *Nat Rev Cancer* **17**, 79-
738 92 (2017).

739
740 34. Smeets D*, et al.* Copy number load predicts outcome of metastatic colorectal cancer patients
741 receiving bevacizumab combination therapy. *Nat Commun* **9**, 4112 (2018).

742
743 35. Gillespie M*, et al.* The reactome pathway knowledgebase 2022. *Nucleic Acids Res* **50**, D687-
744 D692 (2022).

745
746 36. Szklarczyk D*, et al.* The STRING database in 2023: protein-protein association networks and
747 functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* **51**,
748 D638-D646 (2023).

749
750 37. Schubert M*, et al.* Perturbation-response genes reveal signaling footprints in cancer gene
751 expression. *Nat Commun* **9**, 20 (2018).

752
753 38. Zhang D*, et al.* ASCL2 induces an immune excluded microenvironment by activating cancer-
754 associated fibroblasts in microsatellite stable colorectal cancer. *Oncogene* **42**, 2841-2853
755 (2023).

756
757 39. Nguyen TT*, et al.* Mutations in the IFNgamma-JAK-STAT Pathway Causing Resistance to
758 Immune Checkpoint Inhibitors in Melanoma Increase Sensitivity to Oncolytic Virus
759 Treatment. *Clin Cancer Res* **27**, 3432-3442 (2021).

760
761 40. Coles GL*, et al.* Unbiased Proteomic Profiling Uncovers a Targetable GNAS/PKA/PP2A Axis in
762 Small Cell Lung Cancer Stem Cells. *Cancer Cell* **38**, 129-143 e127 (2020).

763
764 41. Ito Y, Bae SC, Chuang LS. The RUNX family: developmental regulators in cancer. *Nat Rev*
765 *Cancer* **15**, 81-95 (2015).

766

767 42. Mevel R, Draper JE, Lie ALM, Kouskoff V, Lacaud G. RUNX transcription factors: orchestrators
768 of development. *Development* **146**, (2019).

769
770 43. Xu H, *et al.* Transcription factors in colorectal cancer: molecular mechanism and therapeutic
771 implications. *Oncogene* **40**, 1555-1569 (2021).

772
773 44. Subramanian A, *et al.* Gene set enrichment analysis: a knowledge-based approach for
774 interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550
775 (2005).

776
777 45. Sturm G, *et al.* Comprehensive evaluation of transcriptome-based cell-type quantification
778 methods for immuno-oncology. *Bioinformatics* **35**, i436-i445 (2019).

779
780 46. Guinney J, *et al.* The consensus molecular subtypes of colorectal cancer. *Nat Med* **21**, 1350-
781 1356 (2015).

782
783 47. Nitsche U, *et al.* Mucinous and signet-ring cell colorectal cancers differ from classical
784 adenocarcinomas in tumor biology and prognosis. *Ann Surg* **258**, 775-782; discussion 782-773
785 (2013).

786
787 48. Linnekamp JF, *et al.* Consensus molecular subtypes of colorectal cancer are recapitulated in
788 in vitro and in vivo models. *Cell Death Differ* **25**, 616-633 (2018).

789
790 49. Oshima S, *et al.* Interferon regulatory factor 1 (IRF-1) and IRF-2 distinctively up-regulate gene
791 expression and production of interleukin-7 in human intestinal epithelial cells. *Mol Cell Biol*
792 **24**, 6298-6310 (2004).

793
794 50. Liao W, *et al.* KRAS-IRF2 Axis Drives Immune Suppression and Immune Therapy Resistance in
795 Colorectal Cancer. *Cancer Cell* **35**, 559-572 e557 (2019).

796
797 51. Horak P, *et al.* Comprehensive Genomic and Transcriptomic Analysis for Guiding Therapeutic
798 Decisions in Patients with Rare Cancers. *Cancer Discov* **11**, 2780-2795 (2021).

799
800 52. Prahallad A, *et al.* Unresponsiveness of colon cancer to BRAF(V600E) inhibition through
801 feedback activation of EGFR. *Nature* **483**, 100-103 (2012).

802
803 53. Troiani T, *et al.* Primary and acquired resistance of colorectal cancer cells to anti-EGFR
804 antibodies converge on MEK/ERK pathway activation and can be overcome by combined
805 MEK/EGFR inhibition. *Clin Cancer Res* **20**, 3775-3786 (2014).

806
807 54. Kim D, *et al.* Pan-KRAS inhibitor disables oncogenic signalling and tumour growth. *Nature*
808 **619**, 160-166 (2023).

809
810 55. Pahuja KB, *et al.* Actionable Activating Oncogenic ERBB2/HER2 Transmembrane and
811 Juxtamembrane Domain Mutations. *Cancer Cell* **34**, 792-806 e795 (2018).

812
813 56. Wang C, Jette N, Moussienko D, Bebb DG, Lees-Miller SP. ATM-Deficient Colorectal Cancer
814 Cells Are Sensitive to the PARP Inhibitor Olaparib. *Transl Oncol* **10**, 190-196 (2017).

815
816 57. Behl D*, et al.* Olaparib (O) in patients (pts) with colorectal cancer (CRC) with ATM mutation
817 (mut): Results from the Targeted Agent and Profiling Utilization Registry (TAPUR) study.
818 *Journal of Clinical Oncology* **41**, 122-122 (2023).

819
820 58. Neshat MS*, et al.* Enhanced sensitivity of PTEN-deficient tumors to inhibition of FRAP/mTOR.
821 *Proc Natl Acad Sci U S A* **98**, 10314-10319 (2001).

822
823 59. Ma L*, et al.* Identification of S664 TSC2 phosphorylation as a marker for extracellular signal-
824 regulated kinase mediated mTOR activation in tuberous sclerosis and human cancer. *Cancer*
825 *Res* **67**, 7106-7112 (2007).

826
827 60. Fritsche-Guenther R*, et al.* Alterations of mTOR signaling impact metabolic stress resistance
828 in colorectal carcinomas with BRAF and KRAS mutations. *Sci Rep* **8**, 9204 (2018).

829
830 61. Rousseau B*, et al.* The Spectrum of Benefit from Checkpoint Blockade in Hypermutated
831 Tumors. *N Engl J Med* **384**, 1168-1170 (2021).

832
833 62. Westcott PMK*, et al.* Mismatch repair deficiency is not sufficient to elicit tumor
834 immunogenicity. *Nat Genet* **55**, 1686-1695 (2023).

835
836 63. Becker WR*, et al.* Single-cell analyses define a continuum of cell state and composition
837 changes in the malignant transformation of polyps to colorectal cancer. *Nat Genet* **54**, 985-
838 995 (2022).

839
840 64. Li Q*, et al.* RUNX1 promotes tumour metastasis by activating the Wnt/beta-catenin signalling
841 pathway and EMT in colorectal cancer. *J Exp Clin Cancer Res* **38**, 334 (2019).

842
843 65. Pant S*, et al.* Erdafitinib in patients with advanced solid tumours with FGFR alterations
844 (RAGNAR): an international, single-arm, phase 2 study. *Lancet Oncol* **24**, 925-935 (2023).

845
846 66. Jung B, Staudacher JJ, Beauchamp D. Transforming Growth Factor beta Superfamily Signaling
847 in Development of Colorectal Cancer. *Gastroenterology* **152**, 36-52 (2017).

848
849 67. Xue Z*, et al.* MAP3K1 and MAP2K4 mutations are associated with sensitivity to MEK inhibitors
850 in multiple cancer models. *Cell Res* **28**, 719-729 (2018).

851
852 68. Ried T*, et al.* The landscape of genomic copy number alterations in colorectal cancer and
853 their consequences on gene expression levels and disease outcome. *Mol Aspects Med* **69**, 48-
854 61 (2019).

855
856 69. Lord CJ, Ashworth A. BRCAness revisited. *Nat Rev Cancer* **16**, 110-120 (2016).

857
858 70. Mestrallet G, Brown M, Bozkus CC, Bhardwaj N. Immune escape and resistance to
859 immunotherapy in mismatch repair deficient tumors. *Front Immunol* **14**, 1210164 (2023).

860
861 71. Guo Y, *et al.* Opposing roles of ZEB1 in the cytoplasm and nucleus control cytoskeletal
862 assembly and YAP1 activity. *Cell Rep* **41**, 111452 (2022).

863
864 72. Huang Y, Hong W, Wei X. The molecular mechanisms and therapeutic strategies of EMT in
865 tumor progression and metastasis. *J Hematol Oncol* **15**, 129 (2022).

866
867 73. Lal N, *et al.* KRAS Mutation and Consensus Molecular Subtypes 2 and 3 Are Independently
868 Associated with Reduced Immune Infiltration and Reactivity in Colorectal Cancer. *Clin Cancer*
869 *Res* **24**, 224-233 (2018).

870
871 74. Garralda E, *et al.* MYC targeting by OMO-103 in solid tumors: a phase 1 trial. *Nat Med* **30**,
872 762-771 (2024).

873
874 75. Li X, Kaur N, Albahrani M, Karpf AR, Black AR, Black JD. Crosstalk between protein kinase C
875 alpha and transforming growth factor beta signaling mediated by Runx2 in intestinal
876 epithelial cells. *J Biol Chem* **299**, 103017 (2023).

877
878 76. Yan X, *et al.* RUNX2 interacts with BRG1 to target CD44 for promoting invasion and migration
879 of colorectal cancer cells. *Cancer Cell Int* **20**, 505 (2020).

880
881 77. Tuo Z, *et al.* RUNX1 is a promising prognostic biomarker and related to immune infiltrates of
882 cancer-associated fibroblasts in human cancers. *BMC Cancer* **22**, 523 (2022).

883
884 78. Heide T, *et al.* The co-evolution of the genome and epigenome in colorectal cancer. *Nature*
885 **611**, 733-743 (2022).

886
887 79. Joanito I, *et al.* Single-cell and bulk transcriptome sequencing identifies two epithelial tumor
888 cell states and refines the consensus molecular classification of colorectal cancer. *Nat Genet*
889 **54**, 963-975 (2022).

890
891 80. Khaliq AM, *et al.* Refining colorectal cancer classification and clinical stratification through a
892 single-cell atlas. *Genome Biol* **23**, 113 (2022).

893
894 81. Mateo L, *et al.* Personalized cancer therapy prioritization based on driver alteration co-
895 occurrence patterns. *Genome Med* **12**, 78 (2020).

896
897 82. Rodon J, *et al.* Genomic and transcriptomic profiling expands precision cancer medicine: the
898 WINTHER trial. *Nat Med* **25**, 751-758 (2019).

899
900 83. Wahjudi LW, *et al.* Integrating proteomics into precision oncology. *Int J Cancer* **148**, 1438-
901 1451 (2021).

902
903 84. Reisinger E, *et al.* OTP: An automatized system for managing and processing NGS data. *J*
904 *Biotechnol* **261**, 53-62 (2017).

905
906 85. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-
907 throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010).

908
909 86. Rimmer A, *et al.* Integrating mapping-, assembly- and haplotype-based approaches for calling
910 variants in clinical sequencing applications. *Nat Genet* **46**, 912-918 (2014).

911
912 87. Pagel KA, *et al.* Integrated Informatics Analysis of Cancer-Related Variants. *JCO Clin Cancer*
913 *Inform* **4**, 310-317 (2020).

914
915 88. Wagner AH, *et al.* A harmonized meta-knowledgebase of clinical interpretations of somatic
916 genomic variants in cancer. *Nat Genet* **52**, 448-457 (2020).

917
918 89. Froyen G, *et al.* Standardization of Somatic Variant Classifications in Solid and Haematological
919 Tumours by a Two-Level Approach of Biological and Clinical Classes: An Initiative of the
920 Belgian ComPerMed Expert Panel. *Cancers (Basel)* **11**, (2019).

921
922 90. Morales J, *et al.* A joint NCBI and EMBL-EBI transcript set for clinical genomics and research.
923 *Nature* **604**, 310-315 (2022).

924
925 91. Cunningham F, Moore B, Ruiz-Schultz N, Ritchie GR, Eilbeck K. Improving the Sequence
926 Ontology terminology for genomic variant annotation. *J Biomed Semantics* **6**, 32 (2015).

927
928 92. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection
929 and Visualization from Targeted DNA Sequencing. *PLoS Comput Biol* **12**, e1004873 (2016).

930
931 93. Kleinheinz K, *et al.* ACEseq – allele specific copy number estimation from whole genome
932 sequencing. *bioRxiv*, 210807 (2017).

933
934 94. Olshen AB, Bengtsson H, Neuvial P, Spellman PT, Olshen RA, Seshan VE. Parent-specific copy
935 number in paired tumor-normal studies using circular binary segmentation. *Bioinformatics*
936 **27**, 2038-2046 (2011).

937
938 95. Dobin A, *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

939
940 96. DeLuca DS, *et al.* RNA-SeQC: RNA-seq metrics for quality control and process optimization.
941 *Bioinformatics* **28**, 1530-1532 (2012).

942
943 97. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning
944 sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).

945

946 98. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable read mapping by
947 seed-and-vote. *Nucleic Acids Res* **41**, e108 (2013).

948
949 99. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential
950 expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).

951
952 100. Zhang B*, et al.* Proteogenomic characterization of human colon and rectal cancer. *Nature*
953 **513**, 382-387 (2014).

954
955 101. Vasaikar S*, et al.* Proteogenomic Analysis of Human Colon Cancer Reveals New Therapeutic
956 Opportunities. *Cell* **177**, 1035-1049 e1019 (2019).

957
958 102. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive
959 analysis of somatic variants in cancer. *Genome Res* **28**, 1747-1756 (2018).

960
961 103. Kautto EA*, et al.* Performance evaluation for rapid detection of pan-cancer microsatellite
962 instability with MANTIS. *Oncotarget* **8**, 7452-7463 (2017).

963
964 104. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes
965 among gene clusters. *OMICS* **16**, 284-287 (2012).

966
967 105. Yu G, He QY. ReactomePA: an R/Bioconductor package for reactome pathway analysis and
968 visualization. *Mol Biosyst* **12**, 477-479 (2016).

969
970 106. Barbie DA*, et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers
971 require TBK1. *Nature* **462**, 108-112 (2009).

972
973 107. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular
974 Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425 (2015).

975
976 108. Ritchie ME*, et al.* limma powers differential expression analyses for RNA-sequencing and
977 microarray studies. *Nucleic Acids Res* **43**, e47 (2015).

978
979 109. Holland CH*, et al.* Robustness and applicability of transcription factor and pathway analysis
980 tools on single-cell RNA-seq data. *Genome Biol* **21**, 36 (2020).

981
982 110. Badia IMP*, et al.* decoupleR: ensemble of computational methods to infer biological activities
983 from omics data. *Bioinform Adv* **2**, vbac016 (2022).

984
985 111. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer New York, NY (2009).

986
987 112. Krug K*, et al.* A Curated Resource for Phosphosite-specific Signature Analysis. *Mol Cell*
988 *Proteomics* **18**, 576-593 (2019).

989

990 113. Turei D, *et al.* Integrated intra- and intercellular signaling knowledge for multicellular omics
991 analysis. *Mol Syst Biol* **17**, e9923 (2021).

992
993 114. Lê S, Josse J, Husson F. FactoMineR: AnRPackage for Multivariate Analysis. *Journal of*
994 *Statistical Software* **25**, 1 - 18 (2008).

995
996 115. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in
997 multidimensional genomic data. *Bioinformatics* **32**, 2847-2849 (2016).

998
999 116. Manders F, *et al.* MutationalPatterns: the one stop shop for the analysis of mutational
1000 processes. *BMC Genomics* **23**, 134 (2022).

1001
1002 117. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. *Stat*
1003 *Med* **9**, 811-818 (1990).

1004
1005 118. Boyle EI, *et al.* GO::TermFinder--open source software for accessing Gene Ontology
1006 information and finding significantly enriched Gene Ontology terms associated with a list of
1007 genes. *Bioinformatics* **20**, 3710-3715 (2004).

1008
1009 119. Horak P, *et al.* Assigning evidence to actionability: An introduction to variant interpretation in
1010 precision cancer medicine. *Genes Chromosomes Cancer* **61**, 303-313 (2022).

1011

1012

1013 Acknowledgements

1014 We thank the teams of the DKFZ High-Throughput Sequencing and the OMICS IT Core Facilities for
1015 performing excellent services. We thank Julio Saez-Rodriguez and Aurelien Dugourd (Heidelberg
1016 University), and Marcos Díaz-Gay (University of California San Diego) for support concerning the
1017 implementation of pathway/TF activities, and COSMIC signature analysis, respectively. This study was
1018 supported by the Helmholtz European Partnering Program "Athens Comprehensive Cancer Center
1019 (ACCC)" in the course of a strategic collaboration between the National Hellenic Research Foundation
1020 (NHRF), the G. Gennimatas Hospital, the Medical School of the University of Athens, and the German
1021 Cancer Research Center (DKFZ).

1022 Author contributions

1023 G.N.Z. **clinically assessed or operated patients and coordinated sample acquisition**. S.T., G.N.Z., V.K.,
1024 S.R, Ko.P, and A.P. **collected and processed specimens**. S.T., S.Wo and S.Wi. **oversaw sequencing of**
1025 **the samples**. E-I.V., K.V., G.W., R.K., Ky.P., and O.P. **carried out the bioinformatics analyses and**
1026 **processed data**. E-I.V., K.V., J.B., O.P. and S.Wi. **analyzed and interpreted data, and wrote the**
1027 **manuscript**. E-I.V., A.D-S., G.N.Z., A.P. and S.Wi. **conceived the study**. **All authors participated in**
1028 **discussions and interpretation of the data and results.**

1029 ## Funding

1030 Open Access funding enabled and organized by Project DEAL.

1031 ## Competing interests

1032 The authors declare no competing interests.

1033 ## Additional Information

1034 Correspondence should be addressed to Efstathios-Iason Vlachavas, Olga Papadodima or Stefan
1035 Wiemann.

1036

1037 ## Figure Legends

1038 **Fig. 1. Mutational landscape in the tumors of the ACCC CRC cohort.** Oncoplot presenting scored
1039 genes common with the IntoGen database for 27 ACCC-CRC patients, which were grouped into MSI
1040 and MSS subtypes (MSI_Status). Each column represents a profiled patient and each row a gene of
1041 interest. The alteration frequencies of 54 genes (SNVs, InDels and CNVs) are shown for each patient
1042 and for each gene on the top and on the right barplots, respectively. Genes are sorted based on their
1043 total mutation frequency (depicted on the left). Colored cells indicate respective types of alterations
1044 explained on the right, while gray cells indicate lack of mutations in respective tumors. The
1045 mutational spectrum for each patient showing the relative contribution of each of the six base

1046 substitution types is presented below the top bar plot as a stacked bar chart, using the Mutational
1047 Patterns R package [116] (version 3.4.1).

1048 **Fig. 2**. **Lollipop plots of REACTOME pathways significantly enriched in MSI vs. MSS tumors of the**
1049 **TCGA and CPTAC cohorts.** Functional enrichment analysis was performed using clusterProfiler [104]
1050 (v.4.2.2), and ggplot2 (v.3.4.3) was used for the visualization of enrichment results. P-values were
1051 adjusted for multiple comparisons (p.adjust) [117]. Sizes of bullets represent the number of over-
1052 represented genes (Count) in the corresponding pathway and adjusted p-values (p.adjust) are color-
1053 coded. The Rich Factor value (x-axis) describes the relative ratio of the enriched genes in each
1054 pathway, divided by the total number of annotated genes in the respective set (i.e., over-
1055 representation analysis) [118].

1056 **Fig. 3**. **Significantly differentially activated pathways (A) and transcription factors (TFs) (B) between**
1057 **MSI and MSS CRC tumors in three CRC cohorts (CPTAC, TCGA and ACCC).** Pathway and TF relative
1058 activities were estimated using PROGENy [37] and DorothEA [109] R packages, respectively. Barplot
1059 visualizes normalized Enrichment Scores (NES) with three different shades of red and blue color
1060 denoting significant pathway/TF activity levels of MSI and MSS tumors in CPTAC, TCGA, and ACCC
1061 CRC datasets, respectively. Horizontal dotted lines at +2/-2 values indicate the lower cut-off values
1062 for statistically significant activities of pathways (A) and TFs (B). Only pathways and TFs found
1063 significantly differentially activated in all three datasets are illustrated. The total numbers of samples
1064 characterizing each CRC group (MSI or MSS) are denoted in parentheses.

1065 **Fig. 4. Clustering of ACCC-CRC patients based on integrative transcriptomic analysis**. (A) Normalized
1066 enrichment scores (NES) of 7 pathway and 14 TF activities (compare Fig. 3) were used for
1067 unsupervised clustering on principal components (HCPC) with R packages FactoMineR [114] (version
1068 2.8) and Factoshiny (http://factominer.free.fr/graphs/factoshiny.html, version 2.4), distributing ACCC
1069 patients into 3 patient clusters. Features were re-ordered within each Pathway and TF row sub-
1070 clusters using hierarchical clustering ("complete" method, "euclidean" distance), where samples
1071 were firstly grouped based on their cluster ID and subsequently based on their MSI status and

1072 mutational information. Tumors with co-occurring alterations in either in *RAS* and *RAF* or *GNAS* and
1073 *RAF* genes were assigned to the RAS_RAF_GNAS_mut group. Pathway and TF activities (NES) are
1074 indicated in shades of red (high) vs. blue (low). Asterisks in individual cells denote a significant
1075 activation of the respective pathway/TF in a patient (NES ≥2 or NES ≤-2). (B) Heatmap visualization of
1076 the normalized enrichment scores of hallmarks from MsigDb [107] per patient using ssGSEA [44]. Re-
1077 ordering of rows was performed as described in (A). Hallmark pathway activities as well as immune
1078 or stromal scores (ImmDec: immune deconvolution), representing immune and stromal cell
1079 infiltration in respective tumors, are indicated in shades of red (high) vs. blue (low). Plots were
1080 created using the ComplexHeatmap R package [115].

1081 **Fig. 5. SVRACAS prioritized somatic mutations, pathway/TF activities, and immune profiles stratify**
1082 **ACCC CRC patients to MASTER baskets and personalized therapeutic interventions.** (A). Oncoplot
1083 with actionable variants (SVRACAS score ≥0.5) and annotated in the context of sensitivity
1084 (OncoKB_Sens 1-4) or resistance (OncoKB_Res R1-2) to a targeted therapy
1085 (https://www.oncokb.org/levels) in any tumor entity. Columns correspond to patients and rows
1086 denote genes and protein-level types of alterations. Types of alterations are indicated in color, while
1087 wild type condition is indicated in gray. Rows were sorted by the percentage of occurring alterations,
1088 whereas columns/samples were firstly grouped based on patient cluster ID, then by microsatellite
1089 instability status, and subsequently by selected driver genomic alterations (i.e.,
1090 *KRAS*/*BRAF*/RAS_RAF_WT). TMB: Tumor mutational burden. Tumors with co-occurring alterations in
1091 either in *RAS* and *RAF* or *GNAS* and *RAF* genes were assigned to the RAS_RAF_GNAS_mut group. (B).
1092 Mapping of indicated patients to MASTER interventional baskets (MB_assignment in blue) based on
1093 mutated genes that are associated with respective baskets [119] (RME: RAF–MEK–ERK; PAM: PI3K–AKT–
1094 mTOR; DDR: DNA damage repair; TK: tyrosine kinases; OTH: other; DEV: developmental regulation;
1095 CC: cell cycle). Mapping to the immune evasion (IE) basket was based on the presence of MSI and/or
1096 high TMB. MB_Assignment in white: not assigned. (C). Molecular rationale for suggested
1097 personalized therapeutic interventions for two ACCC_CRC patients. The first series of boxes

1098 (aberrations) illustrates actionable variants (in orange) and Tumor Mutational Burden (TMB) as well
1099 as MSI-related COSMIC SBS signatures [32]. The second series (activities) presents perturbed molecular
1100 pathways, TF-MYC, and stromal and immune signatures, with boxes framed in shades of red or blue
1101 to indicate either high or low activity levels and high or low scores, respectively. The final series
1102 (Baskets & drugs) shows interventional baskets [119], with suggested drugs, based on a comprehensive
1103 analysis of actionable genomic variants, molecular pathway/TF activities, and immune profiles.

1104 Supplementary Table Legends

1105 **Supplementary Table 1: Patient characteristics of the ACCC CRC patient cohort.** Age (patient age at
1106 the time of diagnosis), sex (F: female; M: male), tumor_location, histological_subtype
1107 (adenoCa_int_diff: adeno carcinoma with intermediary differentiation; adenoCa_NOS: adeno
1108 carcinoma not otherwise specified; LVI: lymphovascular invasion; src-mucinous_adenoCa: serrated
1109 mucinous adenocarcinoma; adenoCa_low_diff: adeno carcinoma with low differentiation;
1110 desmin_adenoCa: desmin-positive adeno carcinoma), T (tumor), N (lymph node), and M (distant
1111 metastasis) status, and mutation status in *RAS*, *RAF*, *GNAS* genes is provided (NA: not analyzed).

1112 **Supplementary Table 2. MSI-status (MSI MANTIS [103]) provided with the TCGA dataset and PreMSIm**
1113 **[23] predictions for 434 patients of the TCGA PanCan COAD/READ cohort [11].** [1] MSI-high
1114 (Microsatellite instability); [0] MSI-low/MSS (Microsatellite stability).

1115 **Supplementary Table 3. MSI-status provided with the CPTAC dataset and PreMSIm [23] predictions**
1116 **for 79 patients of the CPTAC CRC cohort [24].** [1] MSI-high (Microsatellite instability); [0] MSI-low/MSS
1117 (Microsatellite stability).

1118 **Supplementary Table 4. Tumor mutational burden, total numbers of somatic mutations, numbers**
1119 **of SVRACAS scored mutations, MSI-status prediction and probability for 28 patients of the ACCC**
1120 **CRC cohort.** Microsatellite status was estimated using PreMSIm [23]. [1] MSI-high (Microsatellite
1121 instability); [0] MSI-low/MSS (Microsatellite stability).

1122 **Supplementary Table 5. List of somatic mutations in the ACCC patients, with variant annotation**
1123 **information from OpenCRAVAT/OakVar, along with the SVRACAS scoring estimates.** Ref_Base:
1124 Base in reference sequence; Alt_Base: Base in tumor; Selected_Transcript: Genes with alterations
1125 were initially mapped to MANE Select [90] (version 0.9). Genes that did not map in MANE Select were
1126 next assessed for predicted consequences respective mutations have according to the Sequence
1127 Ontology [91]. The aberrant transcript having the highest predicted impact was selected. The longest
1128 transcript was chosen when the same impact was predicted by Sequence Ontology for several
1129 transcripts representing the same gene; cChange: position and consequence of alteration compared
1130 to reference nucleodide sequence; pChange: position and consequence of alteration compared to
1131 reference protein sequence; SVRACAS_Score: Pathogenicity and score assigned to every alteration
1132 using the SVRACAS scoring pipeline (https://github.com/Jasonmbg/Simple.-Variant-Ranking-
1133 Annotation-CAncer-Score).

1134 **Supplementary Table 6. List of somatic mutations in TCGA patients, with variant annotation**
1135 **information from OpenCRAVAT/OakVar, along with the SVRACAS scoring estimates.** Ref_Base:
1136 Base in reference sequence; Alt_Base: Base in tumor; Selected_Transcript: Genes with alterations
1137 were initially mapped to MANE Select [90] (version 0.9). Genes that did not map in MANE Select were
1138 next assessed for predicted consequences respective mutations have according to the Sequence
1139 Ontology [91]. The aberrant transcript having the highest predicted impact was selected. The longest
1140 transcript was chosen when the same impact was predicted by Sequence Ontology for several
1141 transcripts representing the same gene; cChange: position and consequence of alteration compared
1142 to reference nucleodide sequence; pChange: position and consequence of alteration compared to
1143 reference protein sequence; SVRACAS_Score: Pathogenicity and score assigned to every alteration
1144 using the SVRACAS scoring pipeline (https://github.com/Jasonmbg/Simple.-Variant-Ranking-
1145 Annotation-CAncer-Score).

1146 **Supplementary Table 7. List of somatic mutations in CPTAC patients, with variant annotation**
1147 **information from OpenCRAVAT/OakVar, along with the SVRACAS scoring estimates.** Ref_Base:

1148 Base in reference sequence; Alt_Base: Base in tumor; Selected_Transcript: Genes with alterations
1149 were initially mapped to MANE Select [90] (version 0.9). Genes that did not map in MANE Select were
1150 next assessed for predicted consequences respective mutations have according to the Sequence
1151 Ontology [91]. The aberrant transcript having the highest predicted impact was selected. The longest
1152 transcript was chosen when the same impact was predicted by Sequence Ontology for several
1153 transcripts representing the same gene; cChange: position and consequence of alteration compared
1154 to reference nucleodide sequence; pChange: position and consequence of alteration compared to
1155 reference protein sequence; SVRACAS_Score: Pathogenicity and score assigned to every alteration
1156 using the SVRACAS scoring pipeline (https://github.com/Jasonmbg/Simple.-Variant-Ranking-
1157 Annotation-CAncer-Score).

1158 Supplementary Figure Legends

1159 **Supplementary Fig. 1. Relative contribution of COSMIC mutational signatures in each patient of the**
1160 **ACCC-CRC cohort.** Single base substitution (SBS) signatures [32] were assessed using the
1161 SigProfilerAssignment web tool and the ggplot2 R package (v.3.4.3). Signatures accounting for similar
1162 biological processes were grouped together into the same "color" representation in the barplot,
1163 based on the following scheme: MSI-related signatures [SBS6, SBS15, SBS20, SBS26] and POLE-
1164 related mutational procedures [SBS10a and SBS10b].

1165 **Supplementary Fig. 2. K-means clustering of STRING protein-protein interaction network of high**
1166 **mutational frequency genes in MSI versus MSS colorectal cancer.** The network was generated using
1167 the STRING database (version 12.0) with an input of 60 genes identified to have a higher mutational
1168 frequency in MSI compared to MSS colorectal cancer (CRC) tumors. Interactions were considered at a
1169 confidence threshold score ≥0.7. The network was clustered into groups using k-means clustering.
1170 Respective genes within clusters are color coded. Solid lines indicate direct functional or/and physical
1171 protein associations within clusters, while dotted lines represent interactions between clusters. Gene
1172 symbols are indicated in the network.

1173 **Supplementary Fig. 3. Kinase and Pathway Activities, and GNAS Alterations in MSS and MSI CRC**
1174 **Tumors from the CPTAC cohort.** (A) Barplot of the top-10 most activated (red bars) and most
1175 repressed (blue bars) kinases between MSS and MSI colorectal cancer (CRC) tumors, based on
1176 differential phospho-proteomic analysis from the CPTAC patient cohort, with normalized enrichment
1177 scores (NES) from the OmnipathR-decoupleR pipeline [110]. Inferred kinase activities are indicated in
1178 red (higher in MSS, NES ≥2) and blue (higher in MSI, NES ≤-2) colors, respectively. (B) Barplot of
1179 estimated pathway activities, showing top-selected activated (red) and repressed (blue) pathways,
1180 with NES derived from PTM-SEA [112] using PTMSigDB pathway collection (version 2.0.0). (C) Heatmap
1181 of GNAS copy number alterations (from GISTIC 2.0) alongside scaled gene expression values (RNAseq
1182 data) of CPTAC-CRC tumors, stratified by their MSI status. (D) Violin plot detailing GNAS normalized
1183 expression values between MSS and MSI tumors of CPTAC cohort.

Figure 1

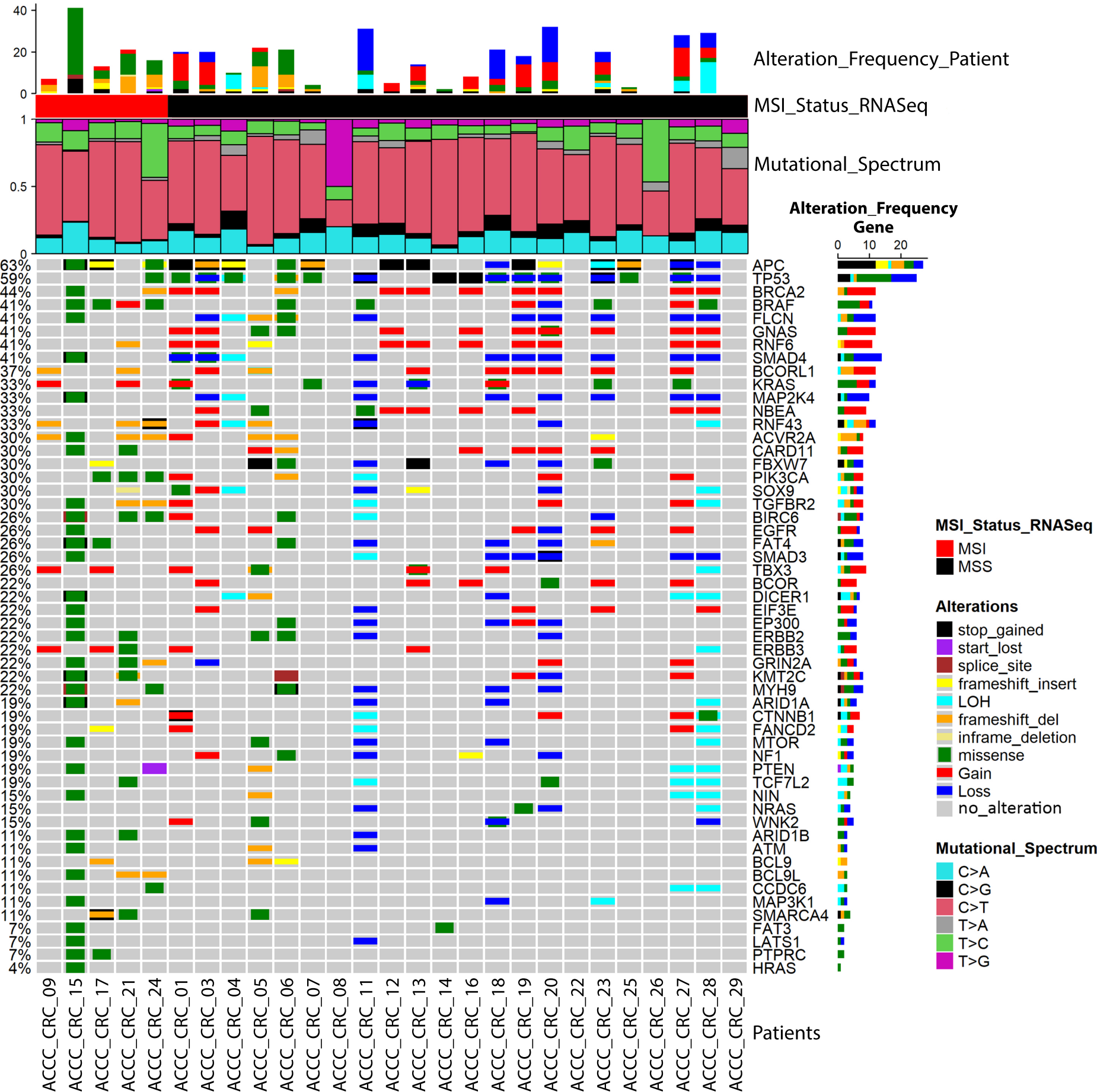

Alteration_Frequency_Patient
MSI_Status_RNASeq
Mutational_Spectrum
Alteration_Frequency
Gene
APC
TP53
BRCA2
BRAF
FLCN
GNAS
RNF6
SMAD4
BCORL1
KRAS
MAP2K4
NBEA
RNF43
ACVR2A
CARD11
FBXW7
PIK3CA
SOX9
TGFBR2
BIRC6
EGFR
FAT4
SMAD3
TBX3
BCOR
DICER1
EIF3E
EP300
ERBB2
ERBB3
GRIN2A
KMT2C
MYH9
ARID1A
CTNNB1
FANCD2
MTOR
NF1
PTEN
TCF7L2
NIN
NRAS
WNK2
ARID1B
ATM
BCL9
BCL9L
CCDC6
MAP3K1
SMARCA4
FAT3
LATS1
PTPRC
HRAS
Patients
MSI_Status_RNASeq
MSI
MSS
Alterations
stop_gained
start_lost
splice_site
frameshift_insert
LOH
frameshift_del
inframe_deletion
missense
Gain
Loss
no_alteration
Mutational_Spectrum
C>A
C>G
C>T
T>A
T>C
T>G

Figure 2

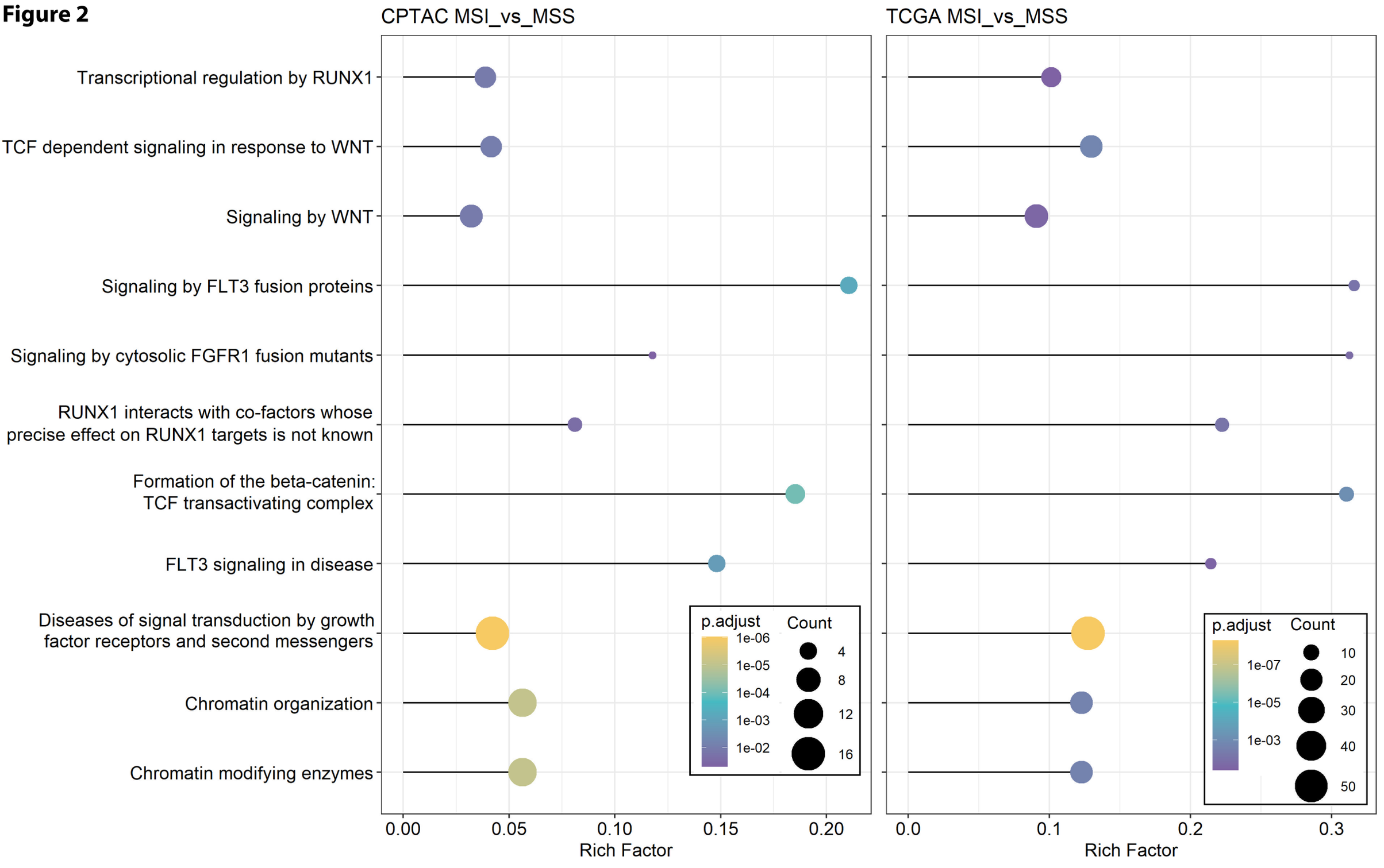

CPTAC MSI_vs_MSS
TCGA MSI_vs_MSS
Transcriptional regulation by RUNX1
TCF dependent signaling in response to WNT
Signaling by WNT
Signaling by FLT3 fusion proteins
Signaling by cytosolic FGFR1 fusion mutants
RUNX1 interacts with co-factors whose precise effect on RUNX1 targets is not known
Formation of the beta-catenin: TCF transactivating complex
FLT3 signaling in disease
Diseases of signal transduction by growth factor receptors and second messengers
Chromatin organization
Chromatin modifying enzymes
0.00
0.05
0.10
0.15
0.20
Rich Factor
0.0
0.1
0.2
0.3
Rich Factor
p.adjust
1e-06
1e-05
1e-04
1e-03
1e-02
Count
4
8
12
16
p.adjust
1e-07
1e-05
1e-03
Count
10
20
30
40
50

Figure 3

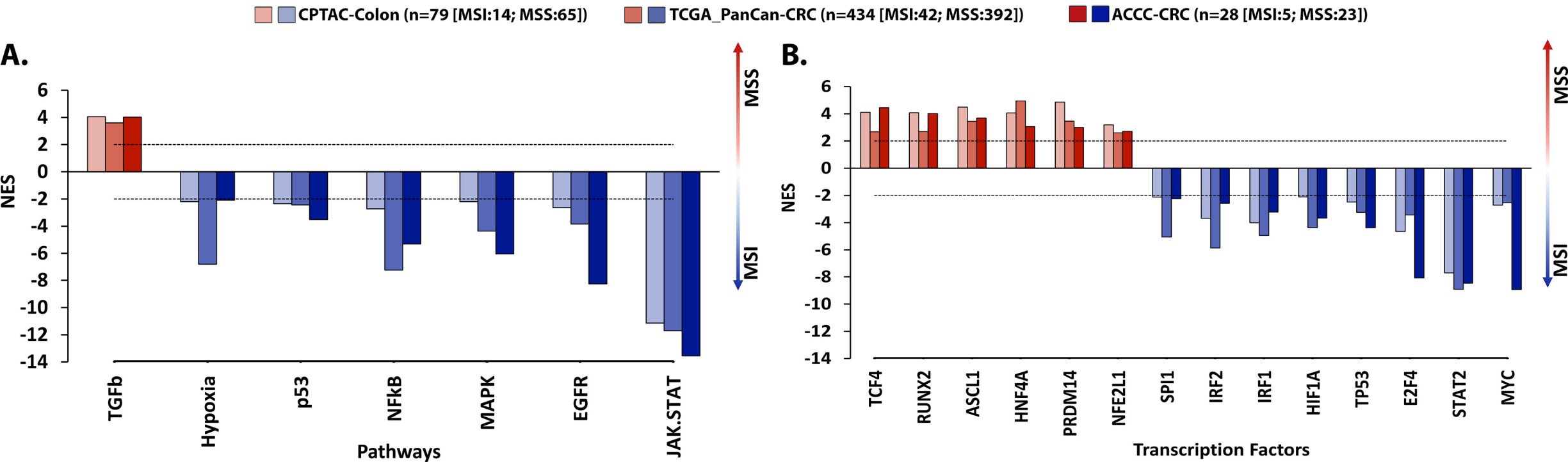

CPTAC-Colon (n=79 [MSI:14; MSS:65])
TCGA_PanCan-CRC (n=434 [MSI:42; MSS:392])
ACCC-CRC (n=28 [MSI:5; MSS:23])
A.
NES
6
4
2
0
-2
-4
-6
-8
-10
-12
-14
MSS
MSI
TGFb
Hypoxia
p53
NFkB
MAPK
EGFR
JAK.STAT
Pathways
B.
NES
6
4
2
0
-2
-4
-6
-8
-10
-12
-14
MSS
MSI
TCF4
RUNX2
ASCL1
HNF4A
PRDM14
NFE2L1
SPI1
IRF2
IRF1
HIF1A
TP53
E2F4
STAT2
MYC
Transcription Factors

Figure 4

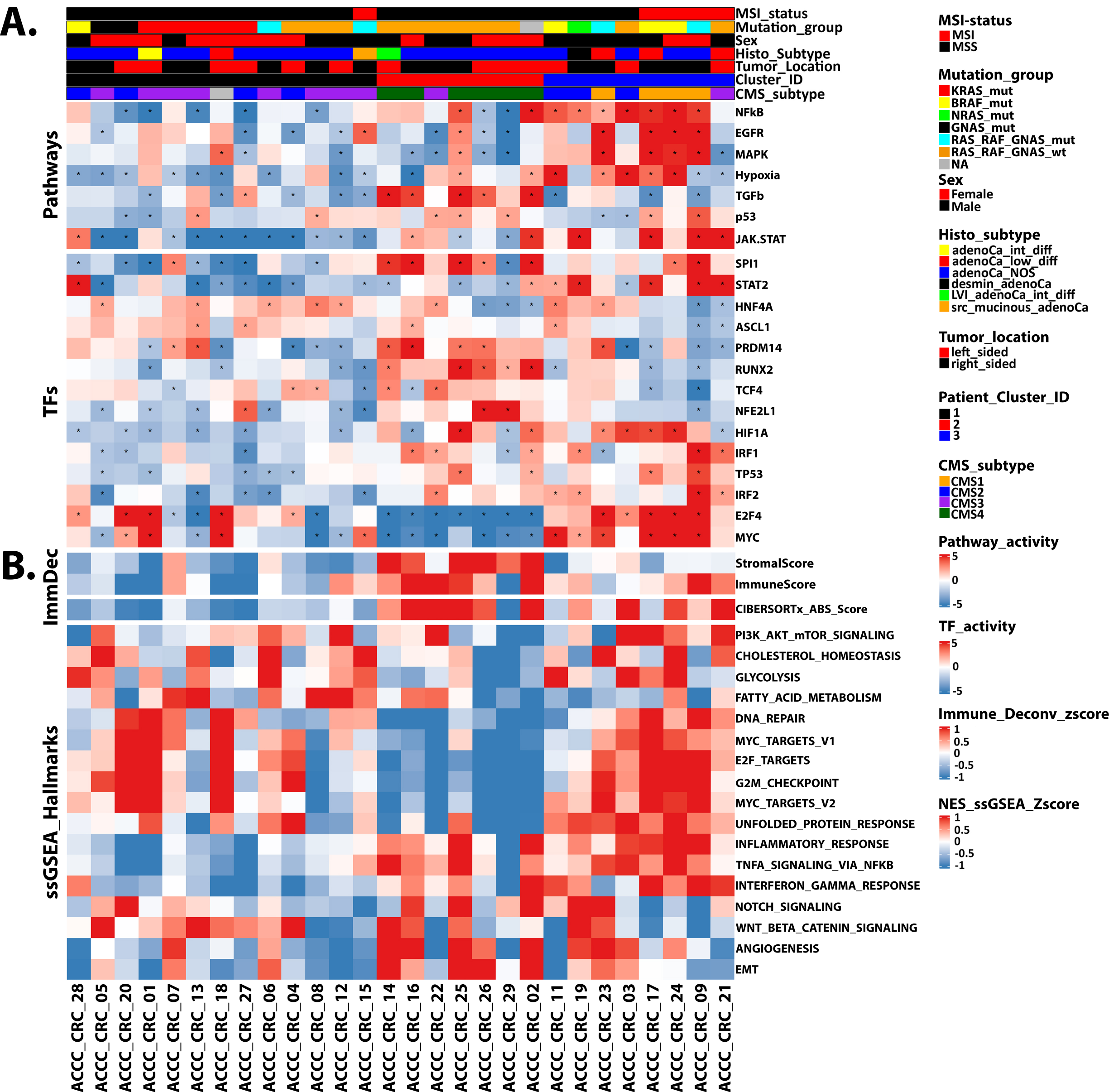
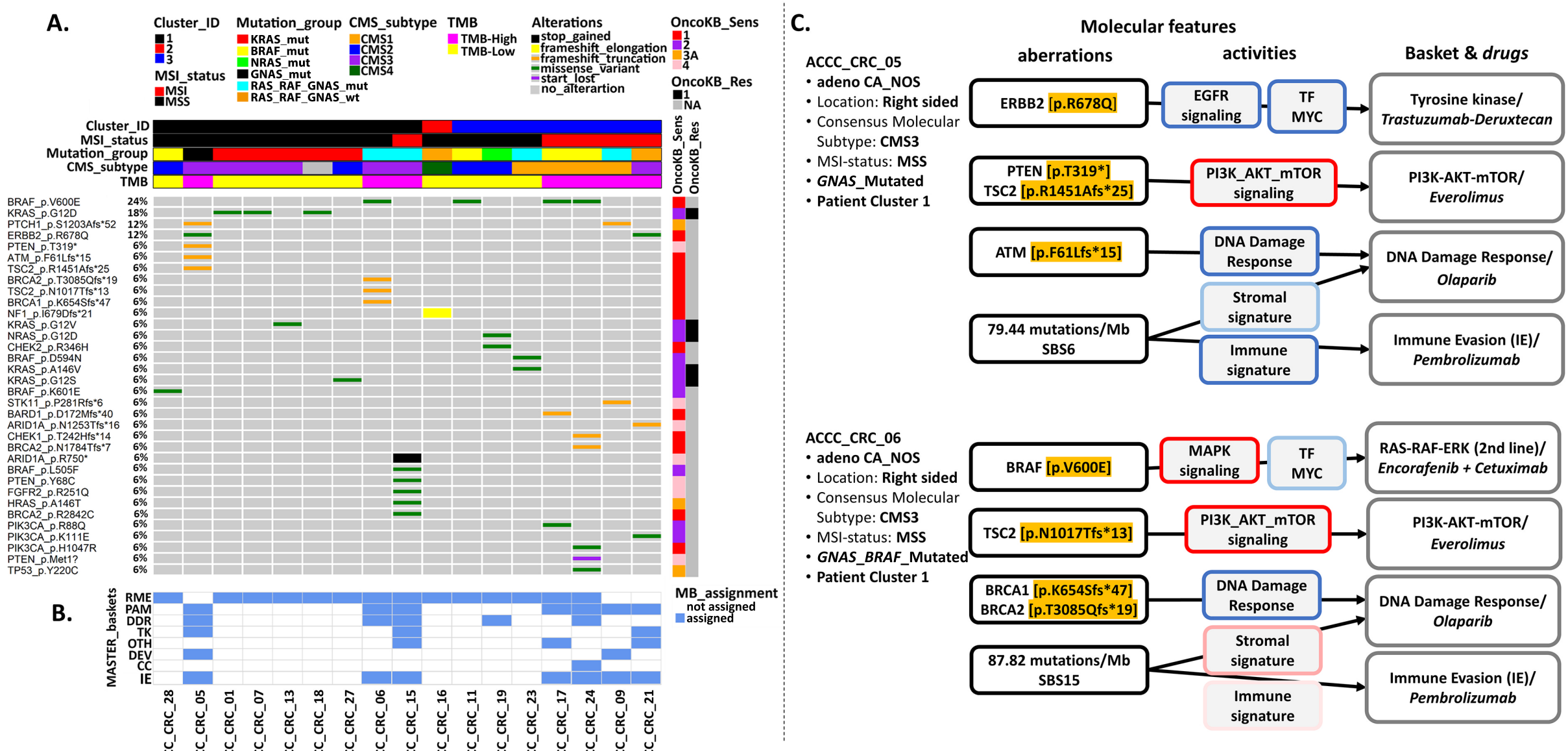

A.
B.
MSI_status
Mutation_group
Sex
Histo_Subtype
Tumor_Location
Cluster_ID
CMS_subtype
Pathways
NFkB
EGFR
MAPK
Hypoxia
TGFb
p53
JAK.STAT
TFs
SPI1
STAT2
HNF4A
ASCL1
PRDM14
RUNX2
TCF4
NFE2L1
HIF1A
IRF1
TP53
IRF2
E2F4
MYC
ImmDec
StromalScore
ImmuneScore
CIBERSORTx_ABS_Score
ssGSEA_Hallmarks
PI3K_AKT_mTOR_SIGNALING
CHOLESTEROL_HOMEOSTASIS
GLYCOLYSIS
FATTY_ACID_METABOLISM
DNA_REPAIR
MYC_TARGETS_V1
E2F_TARGETS
G2M_CHECKPOINT
MYC_TARGETS_V2
UNFOLDED_PROTEIN_RESPONSE
INFLAMMATORY_RESPONSE
TNFA_SIGNALING_VIA_NFKB
INTERFERON_GAMMA_RESPONSE
NOTCH_SIGNALING
WNT_BETA_CATENIN_SIGNALING
ANGIOGENESIS
EMT
ACCC_CRC_28
ACCC_CRC_05
ACCC_CRC_20
ACCC_CRC_01
ACCC_CRC_07
ACCC_CRC_13
ACCC_CRC_18
ACCC_CRC_27
ACCC_CRC_06
ACCC_CRC_04
ACCC_CRC_08
ACCC_CRC_12
ACCC_CRC_15
ACCC_CRC_14
ACCC_CRC_16
ACCC_CRC_22
ACCC_CRC_25
ACCC_CRC_26
ACCC_CRC_29
ACCC_CRC_02
ACCC_CRC_11
ACCC_CRC_19
ACCC_CRC_23
ACCC_CRC_03
ACCC_CRC_17
ACCC_CRC_24
ACCC_CRC_09
ACCC_CRC_21
MSI-status
MSI
MSS
Mutation_group
KRAS_mut
BRAF_mut
NRAS_mut
GNAS_mut
RAS_RAF_GNAS_mut
RAS_RAF_GNAS_wt
NA
Sex
Female
Male
Histo_subtype
adenoCa_int_diff
adenoCa_low_diff
adenoCa_NOS
desmin_adenoCa
LVI_adenoCa_int_diff
src_mucinous_adenoCa
Tumor_location
left_sided
right_sided
Patient_Cluster_ID
1
2
3
CMS_subtype
CMS1
CMS2
CMS3
CMS4
Pathway_activity
5
0
-5
TF_activity
5
0
-5
Immune_Deconv_zscore
1
0.5
0
-0.5
-1
NES_ssGSEA_Zscore
1
0.5
0
-0.5
-1

Figure 5


A.
Cluster_ID
1
2
3
MSI_status
MSI
MSS
Mutation_group
KRAS_mut
BRAF_mut
NRAS_mut
GNAS_mut
RAS_RAF_GNAS_mut
RAS_RAF_GNAS_wt
CMS_subtype
CMS1
CMS2
CMS3
CMS4
TMB
TMB-High
TMB-Low
Alterations
stop_gained
frameshift_elongation
frameshift_truncation
missense_variant
start_lost
no_alterartion
OncoKB_Sens
1
2
3A
4
OncoKB_Res
1
NA
Cluster_ID
MSI_status
Mutation_group
CMS_subtype
TMB
OncoKB_Sens
OncoKB_Res
BRAF_p.V600E 24%
KRAS_p.G12D 18%
PTCH1_p.S1203Afs*52 12%
ERBB2_p.R678Q 12%
PTEN_p.T319* 6%
ATM_p.F61Lfs*15 6%
TSC2_p.R1451Afs*25 6%
BRCA2_p.T3085Qfs*19 6%
TSC2_p.N1017Tfs*13 6%
BRCA1_p.K654Sfs*47 6%
NF1_p.I679Dfs*21 6%
KRAS_p.G12V 6%
NRAS_p.G12D 6%
CHEK2_p.R346H 6%
BRAF_p.D594N 6%
KRAS_p.A146V 6%
KRAS_p.G12S 6%
BRAF_p.K601E 6%
STK11_p.P281Rfs*6 6%
BARD1_p.D172Mfs*40 6%
ARID1A_p.N1253Tfs*16 6%
CHEK1_p.T242Hfs*14 6%
BRCA2_p.N1784Tfs*7 6%
ARID1A_p.R750* 6%
BRAF_p.L505F 6%
PTEN_p.Y68C 6%
FGFR2_p.R251Q 6%
HRAS_p.A146T 6%
BRCA2_p.R2842C 6%
PIK3CA_p.R88Q 6%
PIK3CA_p.K111E 6%
PIK3CA_p.H1047R 6%
PTEN_p.Met1? 6%
TP53_p.Y220C 6%
B.
MASTER_baskets
RME
PAM
DDR
TK
OTH
DEV
CC
IE
MB_assignment
not assigned
assigned
ACCC_CRC_28
ACCC_CRC_05
ACCC_CRC_01
ACCC_CRC_07
ACCC_CRC_13
ACCC_CRC_18
ACCC_CRC_27
ACCC_CRC_06
ACCC_CRC_15
ACCC_CRC_16
ACCC_CRC_11
ACCC_CRC_19
ACCC_CRC_23
ACCC_CRC_17
ACCC_CRC_24
ACCC_CRC_09
ACCC_CRC_21
C.
Molecular features
aberrations
activities
Basket & drugs
ACCC_CRC_05
• adeno CA_NOS
• Location: Right sided
• Consensus Molecular Subtype: CMS3
• MSI-status: MSS
• GNAS_Mutated
• Patient Cluster 1
ERBB2 [p.R678Q]
EGFR signaling
TF MYC
Tyrosine kinase/ Trastuzumab-Deruxtecan
PTEN [p.T319*]
TSC2 [p.R1451Afs*25]
PI3K_AKT_mTOR signaling
PI3K-AKT-mTOR/ Everolimus
ATM [p.F61Lfs*15]
DNA Damage Response
DNA Damage Response/ Olaparib
Stromal signature
79.44 mutations/Mb
SBS6
Immune signature
Immune Evasion (IE)/ Pembrolizumab
ACCC_CRC_06
• adeno CA_NOS
• Location: Right sided
• Consensus Molecular Subtype: CMS3
• MSI-status: MSS
• GNAS_BRAF_Mutated
• Patient Cluster 1
BRAF [p.V600E]
MAPK signaling
TF MYC
RAS-RAF-ERK (2nd line)/ Encorafenib + Cetuximab
TSC2 [p.N1017Tfs*13]
PI3K_AKT_mTOR signaling
PI3K-AKT-mTOR/ Everolimus
BRCA1 [p.K654Sfs*47]
BRCA2 [p.T3085Qfs*19]
DNA Damage Response
DNA Damage Response/ Olaparib
Stromal signature
87.82 mutations/Mb
SBS15
Immune signature
Immune Evasion (IE)/ Pembrolizumab