Mod Pathol 36 (2023) 100012

MODERN PATHOLOGY

USCAP UNITED STATES AND CANADIAN ACADEMY OF PATHOLOGY
Creating a Better Pathologist

Journal homepage: https://modernpathology.org/

Research Article

Unique Patterns of Heterogeneous Mismatch Repair Protein Expression in Colorectal Cancer Unveil Different Degrees of Tumor Mutational Burden and Distinct Tumor Microenvironment Features

Enrico Berrino[a,b], Maria Costanza Aquilano[c], Emanuele Valtorta[c], Vito Amodio[a,d], Giovanni Germano[a,d], Marco Gusmini[c], Katiuscia Gizzi[a,e], Elisabetta Fenocchio[a], Anna Sapino[a,b], Silvia Marsoni[f], Andrea Sartore-Bianchi[c,g], Alberto Bardelli[a,d], Salvatore Siena[c,g], Emanuela Bonoldi[c], Caterina Marchiò[a,b,*]

[a] *Candiolo Cancer Institute, FPO-IRCCS, Candiolo, Italy;* [b] *Department of Medical Sciences, University of Turin, Turin, Italy;* [c] *Niguarda Cancer Center, Grande Ospedale Metropolitano Niguarda, Milano, Italy;* [d] *Department of Oncology, University of Turin, Turin, Italy;* [e] *IIGM-Italian Institute for Genomic Medicine, c/o IRCCS, Candiolo (TO), Italy;* [f] *FIRC Institute of Molecular Oncology (IFOM), Milan, Italy;* [g] *Department of Oncology and Hemato-Oncology, Università degli Studi di Milano, Milan, Italy*

ARTICLE INFO

Article history:
Received 12 May 2022
Revised 9 September 2022
Accepted 16 September 2022
Available online 10 January 2023

Keywords:
colorectal cancer
heterogeneity
immunohistochemistry
microsatellite instability
mismatch repair deficiency
tumor-infiltrating lymphocytes

ABSTRACT

Mismatch repair (MMR) protein expression in colorectal cancer (CRC) cells is usually homogeneously retained or lost. Rare lesions may show a heterogeneous pattern of MMR protein expression. We evaluated MMR protein expression (MLH1, MSH2, MSH6, and PMS2) in 200 CRCs, identifying 3 groups with proficient MMR protein expression (MMRp), deficient MMR protein expression (MMRd), and heterogeneous MMR protein expression (MMRh). MMRh tumors were microdissected on the basis of the expression of the heterogeneous marker. DNA was extracted and subjected to targeted sequencing. RNA was purified from bulk tumors of all MMRh cases and in a control series of 15 MMRp and 10 MMRd CRCs and analyzed using the PanCancer IO 360 Panel (NanoString Technologies). Twenty-nine of the 200 cases (14.5%) were MMRd. Nine cases (4.5%) showed a heterogeneous pattern of MMR expression, with 6 tumors harboring concomitant loss of one of the other MMR proteins, thus featuring areas with double loss at immunohistochemistry (IHC) testing (MMRh double-loss cases). Four of the 6 MMRh double-loss cases were suitable for a separate sequence variant analysis of IHC double-negative and IHC single-negative components of the tumor. In all lesions, both components exhibited a high tumor mutation burden (TMB). Nevertheless, a significant increase in TMB in the double-negative components was observed (mean TMB: negative, 70 mut/Mb vs positive, 59 mut/Mb) because of a higher number of subclonal variants compared with the other component. Comparative gene expression analyses among MMRd, MMRp, and MMRh CRCs highlighted differential gene expression patterns and an increased number of tumor-infiltrating lymphocytes in MMRh lesions, which is also characterized by a substantial population of exhausted $CD8^+$ lymphocytes. We describe a unique subgroup of CRCs showing heterogeneous expression of MMR proteins in a background of concomitant loss of one of the other markers.

These authors contributed equally: Emanuela Bonoldi and Caterina Marchiò.

* Corresponding author.
E-mail addresses: caterina.marchio@unito.it; caterina.marchio@ircc.it (C. Marchiò).

https://doi.org/10.1016/j.modpat.2022.100012

Introduction

Loss of expression of mismatch repair (MMR) proteins is strictly associated with the microsatellite instable tumor phenotype.[1-3] The carcinogenic model for MMR-deficient tumors is represented by the Lynch syndrome, an autosomal dominant disease related to *MLH1*, *MSH2*, *MSH6*, and *PMS2* gene variants leading to loss of expression of MMR proteins. This germline condition is associated with the development of tumor in different organs, including colorectal cancers (CRCs) characterized by the typical absence of preneoplastic lesions (also known as "hereditary nonpolyposis CRC").[4] The loss of MMR proteins precludes MLH1-PMS2 and MSH2-MSH6 heterodimerization, thus disrupting the MMR machinery that is devoted to recognize nucleotide mismatches in the newly synthesized strand during the DNA replication phase.[1] Defects in this DNA repair mechanism induce an accumulation of single-base errors and small insertions and deletions, which is directly correlated with an increase in the tumor mutation burden (TMB)[5,6] and with a specific mutagenetic effect on the DNA microsatellite repeated sequences.[1] The shortening or lengthening of these sequences (microsatellite instability [MSI]) is used as a pathognomonic hallmark of MMR deficiency.[7] Patients with MSI CRC typically exhibit less differentiated tumors with a high prevalence of the mucinous phenotype; however, they do better than patients with microsatellite stability (MSS) both in terms of disease-free and overall survival, regardless of the tumor stage.[1,8] Furthermore, the MSI status is a positive predictor of response to immune checkpoint inhibitors (ICIs).[9-11]

In this context, the evaluation of the 4 main MMR proteins by immunohistochemistry (IHC) and/or the evaluation of the microsatellite status allow to stratify patients with CRC. In general, MMR protein expression is homogeneously distributed across CRCs. Nevertheless, heterogeneous patterns can be encountered.[12] MMR protein expression heterogeneity can present either with a clear-cut pattern featuring 2 different and well-compartmentalized clonal populations (one with proficient and the other with defective MMR machinery, respectively) or with a more jigsaw-like pattern leading to a diffuse intermingling of MMR-positive and MMR-negative areas within the same tumoral lesion.[12,13] Both patterns can affect MSI assessment in diagnostic practice. The former pattern has been more extensively studied in terms of frequency and molecular features. For instance, Loupakis et al[14] reported a prevalence of <1% in a cohort of >1800 CRCs analyzed for MMR protein expression by IHC. These authors demonstrated a correlation between this heterogeneity in the protein expression and the microsatellite molecular features: the proficient MMR component of heterogeneous tumors was characterized by MSS, whereas the MMR-deficient area showed a clear-cut MSI status.[14] Interestingly, this heterogeneity has also been occasionally described in other Lynch syndrome–related diseases, such as endometrial cancer.[15] On the other side, little is known on the implications of the second pattern, which in some instances may coexist already with the complete loss of one of the MMR heterodimers.[16] In addition, an extensive molecular characterization of MMR heterogeneous CRC tumors is missing. In this study, we present a molecular characterization of MMR heterogeneous CRCs, including the landscape of their immune microenvironment.

Materials and Methods

Cohort and IHC Analysis

A retrospective evaluation of 200 consecutive CRCs diagnosed between 2018 and 2020 was performed at the Niguarda Cancer Center by evaluating immunohistochemical reactions of the 4 major MMR deficiency markers (MSH2, MSH6, MLH1, and PMS2). Through this analysis, 3 subcohorts were identified: (1) tumors with preserved MMR protein expression, defined as MMR proficient (MMRp) (n = 168, 84%); (2) tumors with homogeneous loss of expression of MMR proteins, defined as MMR deficient (MMRd) (n = 23, 11.5%); and (3) tumors with heterogeneous expression of MMR proteins, defined as heterogeneous MMR (MMRh) (n = 9, 4.5%) (Fig. 1).

Patients included in the study signed the informed consent within the AlfaOmega project (approved by the Ethical Committee of the Niguarda Cancer Center on April 7, 2020, #145-07042020; NCT04120935 and NCT05101382). For the 9 MMRh cases, MMR proteins were retested by IHC with antibodies raised against MLH1 (G168-15, dilution 1:40; BD Biosciences), MSH2 (FE11, dilution 1:50; Calbiochem, Merck), MSH6 (44, dilution 1:200; BD Biosciences), and PMS2 (A16-4, 1:40; BD Biosciences). Two independent pathologists (C.M. and M.C.A.) reviewed the IHC slides, providing the percentage of positive tumor cells showing MMR expression.

All MMRh and a selected cohort of 15 MMRd and 10 MMRp tumors were also evaluated for the tumor-infiltrating lymphocytes (TILs) by considering geographic distribution of the markers (core and front of invasion of the tumor and intraepithelial component) using anti-CD8 (clone C8/144B, ready-to-use; Dako-Agilent) and anti-CD3 (polyclonal, ready-to-use; Dako-Agilent) antibodies on the Dako Omnis instrument (Dako-Agilent).

DNA and RNA Extraction

Hematoxylin and eosin–stained sections corresponding to the 9 MMRh, 15 MMRp, and 10 MMRd cases were reviewed to assess tumor cellularity. For analyses of tumor bulk, six 8-μm-thick sections were cut, mesodissected, and used to purify nucleic acids. DNA was extracted using the GeneRead DNA FFPE (Formalin Fixed Paraffin Embedded) Kit (Qiagen). RNA was extracted using the High Pure FFPET (Formalin Fixed Paraffin Embedded Tissue) RNA Isolation Kit (Roche). For MMRh case 8, we purified DNA from the normal adjacent colon mucosa for comparative sequencing analysis.

For tumor microdissection of MMRh CRCs, a total of 10 8-μm-thick sections were cut and IHC stained with the antibody for the heterogeneous marker and microdissected under a stereomicroscope to select the positive and negative areas. The microdissected cells were then subjected to DNA extraction using the GeneRead DNA FFPE Kit (Qiagen).

DNA and RNA samples were quantified with both spectrophotometric (Nanodrop 1000; ThermoFisher Scientific) and fluorometric (Qubit; ThermoFisher Scientific) assays. The integrity of the RNA was also assessed using the BioAnalyzer 6000 Nano Assay on the BioAnalyzer 2100 instrument (Agilent).

Next-Generation Sequencing Analysis

DNA sequencing was applied to a total of 40 ng of DNA using the TruSight Oncology 500 (Illumina) targeted panel (panel size: 1.94 Mb, encoding 1.2 Mb), which includes 523 cancer-related genes.[17,18] The panel allows the evaluation of the entire coding

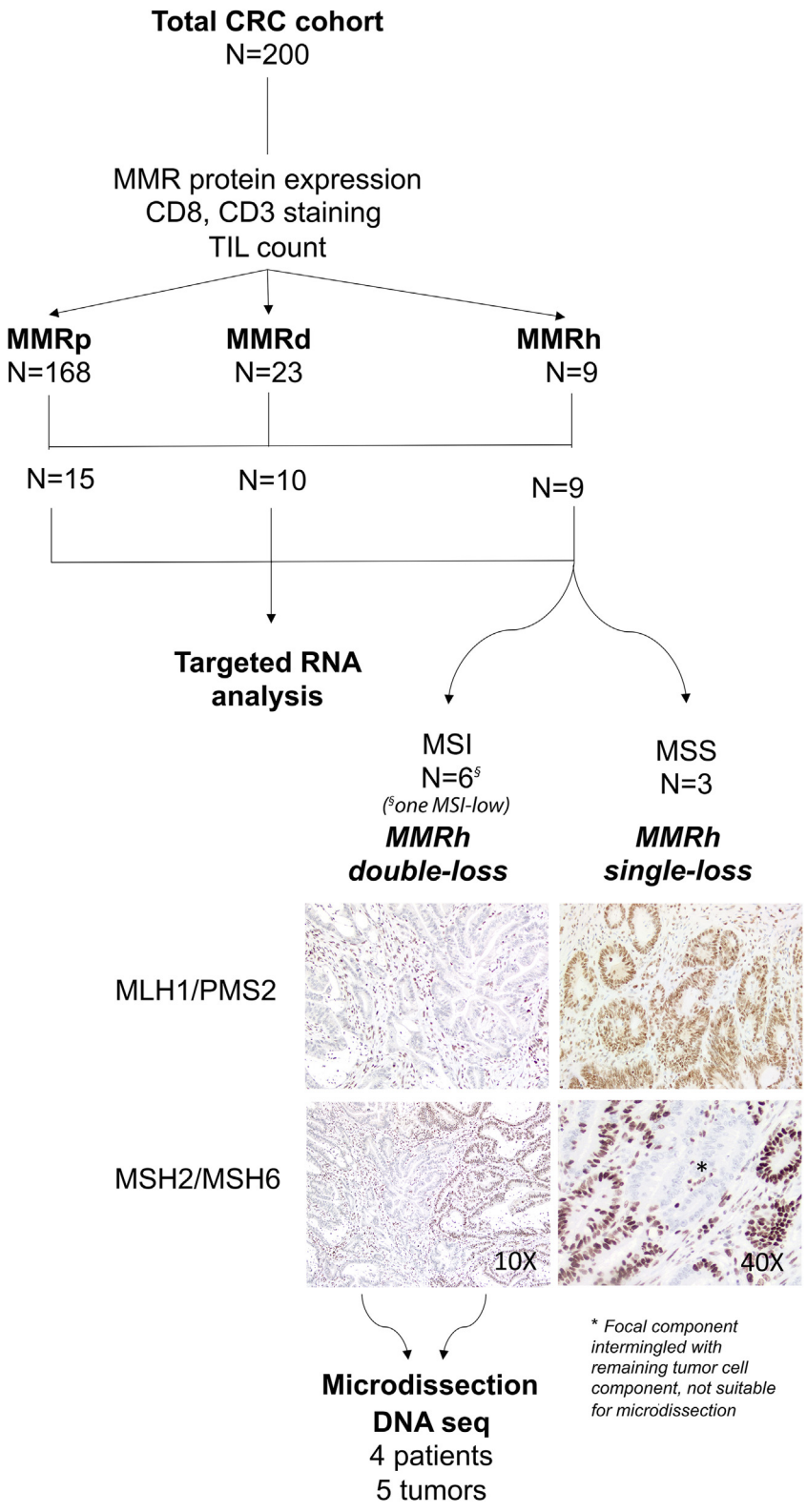

Figure 1.
A description of the analytical cohort. The diagram illustrates the structure of the cohort and the performed analysis. The asterisk indicates the limited heterogeneous areas in the MMRh "single-loss" cases, which can be appreciated only at high magnification (×40), as opposed to the larger areas that were microdissected in the MMRh "double-loss" lesions (×10 magnification in the image). CRC, colorectal cancer; d, deficient; h, heterogeneous; MMR, mismatch repair; MSI, microsatellite instability; MSS, microsatellite stability; p, proficient; TIL, tumor-infiltrating lymphocyte.

sequences and exon-flanking regions of genes related to CRC, including *MLH1*, *MSH2*, *MSH6*, *PMS2*, *APC*, *MUTYH*, *STK11*, *BMPR1A*, *SMAD4*, *POLE*, *POLD1*, *CTNNB1*, *TP53*, *KRAS*, *BRAF*, and *NRAS*. In addition, the MSI status of 120 loci and TMB can be assessed. Following the manufacturer's protocol, the generated libraries were sequenced on the Next-Seq 500 (Illumina) instrument. The Illumina local app associated with the TruSight Oncology 500 was used to perform sequencing and contamination quality checks and generate fastq files through alignment with the human reference sequence GRCh37 (hg19). The Illumina app performed variant calling without a normal-tumor pipeline based on querying different genetic databases and taking into account variant allele frequency (VAF). All samples reached a median $\times$500 of read depth with a uniformity of 95% after both index and unique molecular identifier demultiplexing. Only the variants with at least 50 alternative calls and a VAF of >5% were considered in the analysis. The TMB values were calculated over the real, sequenced region for each sample as the total number of nonsynonymous somatic variants per megabase. The ΔTMB value was calculated as the difference between the TMB value of the negative area and that of the positive area of each MMRh tumor. The VCF (variant call format) files obtained from sequencing were then annotated using ANNOVAR (Annotate Variation) tools for pathogenicity.

For genomic data visualization, the R software (v 4.0.3) was used; for oncoPrint, the function oncoPrint of the Complex Heatmap package was implemented. The heatmap VAF code was built using the heatmap2 function of the gplots package. Venn diagrams were created using the VennDiagram package.

PCR-Based MSI Analysis

All purified DNAs were analyzed for the MSI status using the MSI analysis system, version 1.2 (Promega), which includes 5 mononuclear and monomorphic microsatellite loci. The products were analyzed by capillary electrophoresis using an ABI 3100 Genetic Analyzer (Applied Biosystems). The samples were considered MSI when harboring >40% of unstable loci.

Gene Expression Analysis

Targeted gene expression was performed using the PanCancer IO 360 Panel (NanoString Technologies) on RNA extracted from 15 MMRp and 10 MMRd and the cohort of MMRh (n = 9). Two hundred fifty nanograms of RNA were incubated for 14 hours at 56 °C with fluorescent probes to assess the expression of 770 genes. The nPrep Station (NanoString Technologies) purified the probe-target complexes, and the nCounter instrument (NanoString Technologies) scanned the purified samples. The nSolver software (NanoString Technologies) nSolver Advanced Analyses, based on robust R statistics, examined the results for the primary level of analysis with both the pathway score (calculated as the first principal component of the pathway genes normalized expression) and the cell type score (method described here).[17]

Statistical Analyses

Statistical analyses were performed using the R software (v.4.0.3). Paired *t* test was applied for the TMB value distribution analysis. The analysis of variance test was applied for the evaluation of the $CD8^+$ IHC expression and the comparison of gene expression scores, with Tukey post hoc test for the subsequent paired comparison. The Levene test was applied for the variance analysis. *P* values of <.05 were considered statistically significant.

Results

Prevalence of MMR Heterogeneity in the Selected Cohort

Immunohistochemical analysis of the MMR proteins in the consecutive 200 CRCs identified an MMRd prevalence of 14% (29/200) (Fig. 1). The remaining 171 patients had MMRp. After the accurate evaluation of IHC reactions for MSH2/MSH6/MLH1/PMS2, we identified 9 cases showing heterogeneous expression of MMR proteins (MMRh 4.5%). Of note, in 6 of 9 tumors, at least one of the MMR proteins/heterodimers was not expressed and, at the same time, at least another MMR protein was heterogeneously expressed, thus leading to the areas with "double-negativity" at IHC testing. These tumors harboring the unique feature of showing homogeneous loss of a protein/heterodimer and heterogeneous loss of expression of the other protein/heterodimer were labeled as "MMRh double-loss" because IHC testing showed the areas with double-negativity for MMR proteins/heterodimers in a tumor population that already tested negative for a given MMR protein/heterodimer (Fig. 1). Conversely, the remaining 3 cases displayed retained expression of a protein/heterodimer and heterogeneous, focal loss of expression of the other proteins/heterodimers. Hence, they were labeled as "MMRh single-loss" cases.

To assess the microsatellite status of these tumors, DNAs extracted from the tumor bulk were analyzed using the PentaPlex panel: 5 of 9 heterogeneous cases displayed high MSI (MSI-H), 1 harbored a low level of instability (only 1/5 PentaPlex unstable marker), and the remaining 3 were MSS. Supplementary Table S1 summarizes the characteristics of microsatellite status, reporting also the heterogeneous marker.

We double-checked the clinical and familial history of the patients affected by CRC showing this heterogeneous expression of MMR proteins. All patients were at an elderly age, and none of them were affected by hereditary nonpolyposis CRC (Lynch syndrome).

Molecular Characterization of Tumors Displaying Heterogeneity of MMR Protein Expression

To understand the molecular characteristics of the identified MMRh tumors, we extracted DNA from the distinct areas of 4 samples showing adequate cellularity and feasible microdissection. None of the MMRh single-loss cases could be microdissected because the areas lacking expression were very focal and profoundly intermingled with the remaining tumor cells with retained MMR expression. We were able to microdissect 4 cases that pertained to the "MMRh double-loss" cases and were all MSI-H from bulk analysis (Fig. 1). The MSI status was confirmed for both components in each case. For MMRh case 8 (which was the only MMRh double-loss CRC showing only 1 unstable marker at PCR analysis), although the tumor was not suitable for microdissection, we sequenced the DNA purified from the bulk lesion to confirm the MSI status. We identified 11.1% (10/90) of unstable loci compared with a single altered locus (1/100) in the normal mucosa. In addition, the tumor cells harbored a TMB of 22.4 mut/Mb.

As explained above, MMRh double-loss was composed of areas with double-loss of expression of MMR proteins/heterodimers

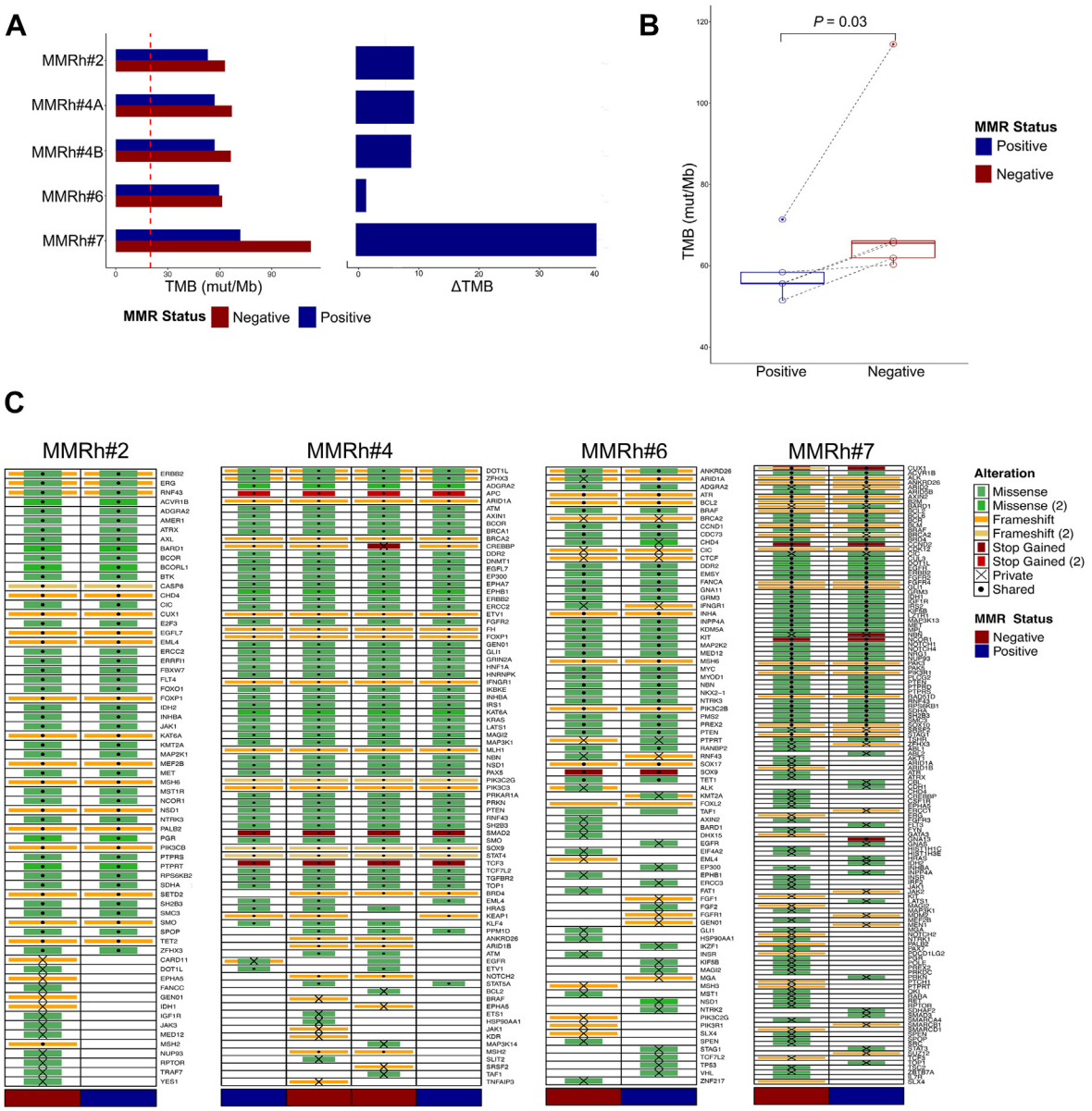

Figure 2.
DNA sequencing results of the 4 MMRh tumors. (A) Bar plots reporting TMB as mutations per megabase and the ΔTMB for each sequenced paired areas. (B) A box plot of paired *t* test, showing the increment of TMB in the double-negative area of the samples. (C) The oncoPrints reporting the identified variants in each pair. h, heterogeneous; MMR, mismatch repair; TMB, tumor mutation burden.

and areas with single-loss of an MMR protein/heterodimer. Of note, microdissection was performed on IHC-stained slides to enable proper identification of the distinct areas. Indeed, on the basis of the pattern of expression of the marker(s) that was heterogeneously distributed across tumors, we could best distinguish those areas with a double-negative IHC pattern from those showing a single-negative IHC pattern.

For MMRh case 4, 2 tissue blocks were available, with a comparable heterogeneity pattern for a total of 10 MMRh samples (5 IHC double-negative areas and 5 IHC single-negative areas) adequate for subsequent molecular analysis. The microdissected DNAs were sequenced using the TruSight Oncology 500 panel. We observed high TMB values for both components of the MMRh tumors (mean TMB positive: 59.7 mut/Mb, range: 53.1-73.8 mut/Mb; mean TMB negative: 74.1 mut/Mb, range: 63-112.6 mut/Mb) (Fig. 2A). Despite this, the tumor regions with a double-negative IHC pattern harbored higher TMB levels: the paired distribution test showed a statistically significant increase in the number of nonsynonymous variants in these tumor areas compared with the remaining areas ($P = .03$, paired *t* test; Fig. 2B). Furthermore, the ΔTMB values highlighted a higher TMB level in areas with a double-negative IHC pattern than that in the other areas (Fig. 2A). This observation was corroborated by the microsatellite analysis derived from the next-generation sequencing panel: a higher number of unstable loci was indeed observed in areas with a double-negative IHC pattern (the percentage of unstable loci in MSI positive is 48.8% and MSI negative is 66.1%) (Supplementary Fig. S1), in line with the output results by the PentaPlex.

The oncoPrints in Figure 2C show the mutated genes in the 8 samples: the areas with double-negative IHC pattern harbored a

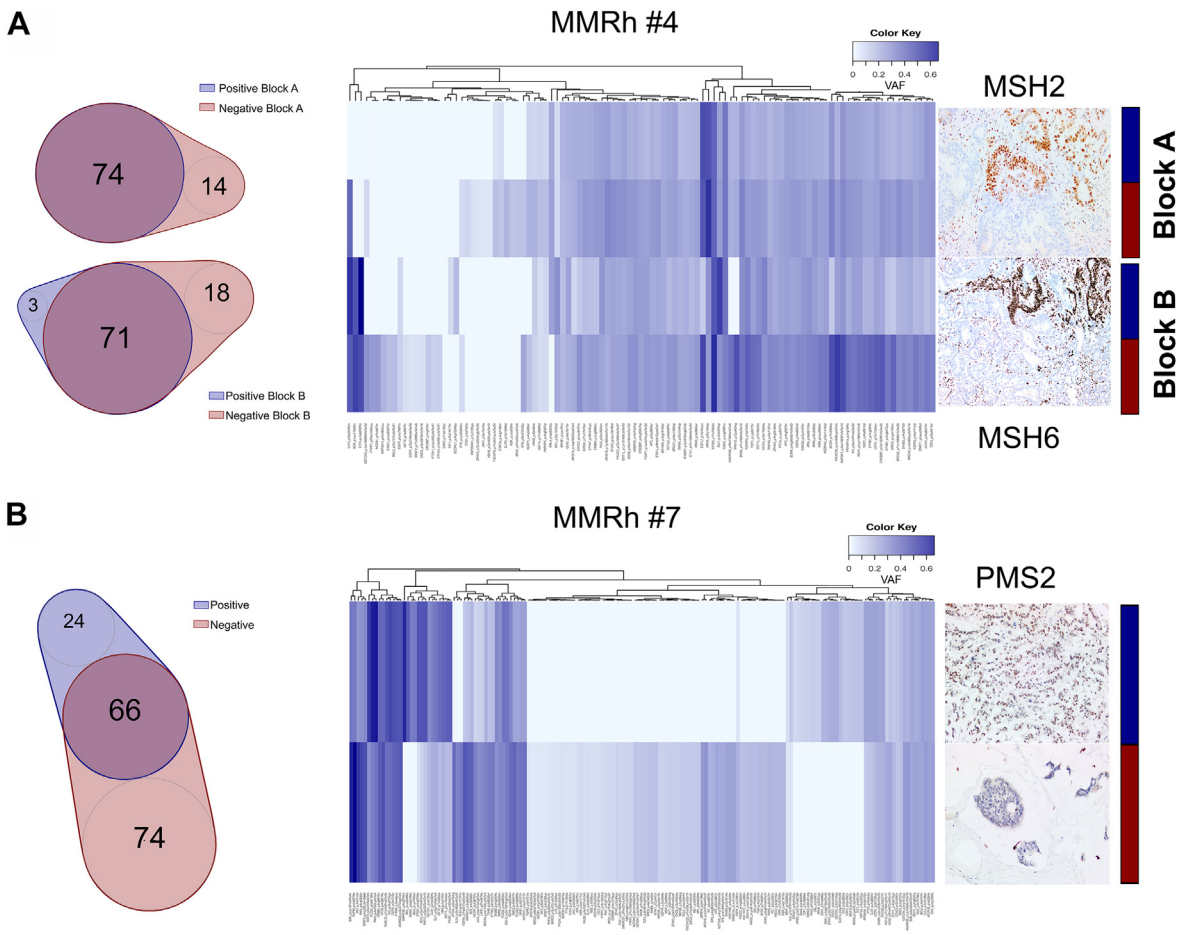

Figure 3.
A landscape of somatic variants of MMRh lesions. (A) A Venn diagram for the private and shared variants; heatmap reporting the VAF and IHC staining of the heterogeneous markers for MMRh case 4. (B) A Venn diagram for the private and shared variants; VAF heatmap reporting and IHC staining of the heterogeneous marker for MMRh case 7. h, heterogeneous; IHC, immunohistochemistry; MMR, mismatch repair; VAF, variant allele frequency.

higher number of variants, which included most of the variants detected in the other areas. All variants are reported in Supplementary Table S2.

To further characterize the molecular features of each tumor, we produced VAF heatmaps and Venn diagrams for the analyzed sample pairs. MMRh case 4 showed heterogeneous MSH2-MSH6 expression associated with the loss of PMS2 expression. As previously specified, 2 different tumor blocks (A and B) were available from this tumor, and microdissection of tumor areas was performed on both blocks with subsequent analysis. This also allowed to assess the reproducibility of results across distinct areas of the same tumor. Of note, the ΔTMB value between the microdissected areas was superimposable in the distinct tumor blocks (10 mut/Mb and 9.5 mut/Mb for inclusion A and B, respectively), with the $MSH2^{+}$-$MSH6^{+}$ regions harboring a lower number of total variants. Almost all private variants belonged to the $MSH2^{-}$-$MSH6^{-}$ fraction (Fig. 3A), with a confirmed subclonal nature ($P < .0001$, unpaired t test; Supplementary Fig. S2).

A comparable scenario was identified in MMRh case 2, which displayed heterogeneous MSH2 expression in a context of complete absence of all other MMR proteins. This tumor showed a strong variant overlap between the 2 microdissected areas. Nevertheless, the $MSH2^{-}$ fraction was characterized by several private subclonal variants ($P < .0001$, unpaired t test; Supplementary Figs. S2 and S3A).

MMRh case 7 was characterized by heterogeneous MLH1-PMS2 expression, with complete loss of the MSH2-MSH6 heterodimer. This tumor harbored the highest ΔTMB value (40.72 mut/Mb) between the areas with a double-negative IHC pattern and those with a single-negative IHC pattern. The tumor bulk showed a marked MSI level using the PentaPlex panel, with Penta C and Penta D markers characterized by the presence of unusual, unstable alleles. This instability corresponded to the highest TMB value among the MMRh tumors (positive area = 71.8 mut/Mb; negative area = 112.6 mut/Mb). The IHC single-negative tumor component of the MMRh case 7 showed several private variants (n = 25). Nevertheless, we confirmed an increased mutagenic trigger in the IHC double-negative tumor component, with a total of 75 exclusive variants (Fig. 3B). The private variants were characterized by a subclonal VAF ($P < .0001$, unpaired t test; Supplementary Fig. S2). Interestingly, the IHC double-negative area showed a *POLE* sequence variation (c.4582G>A, p.Ala1528Thr), defined as "tolerated" or "benign" by pathogenicity tools. This variant has no functional confirmation but has been detected in both hypermutated and low-TMB tumors,[19,20] supporting its "passenger" significance.

The tumor identified as MMRh case 6 displayed heterogeneous MSH6-MSH2 expression with loss of MLH1 expression and exhibited a modest ΔTMB value of 3.25 mut/Mb. The Venn

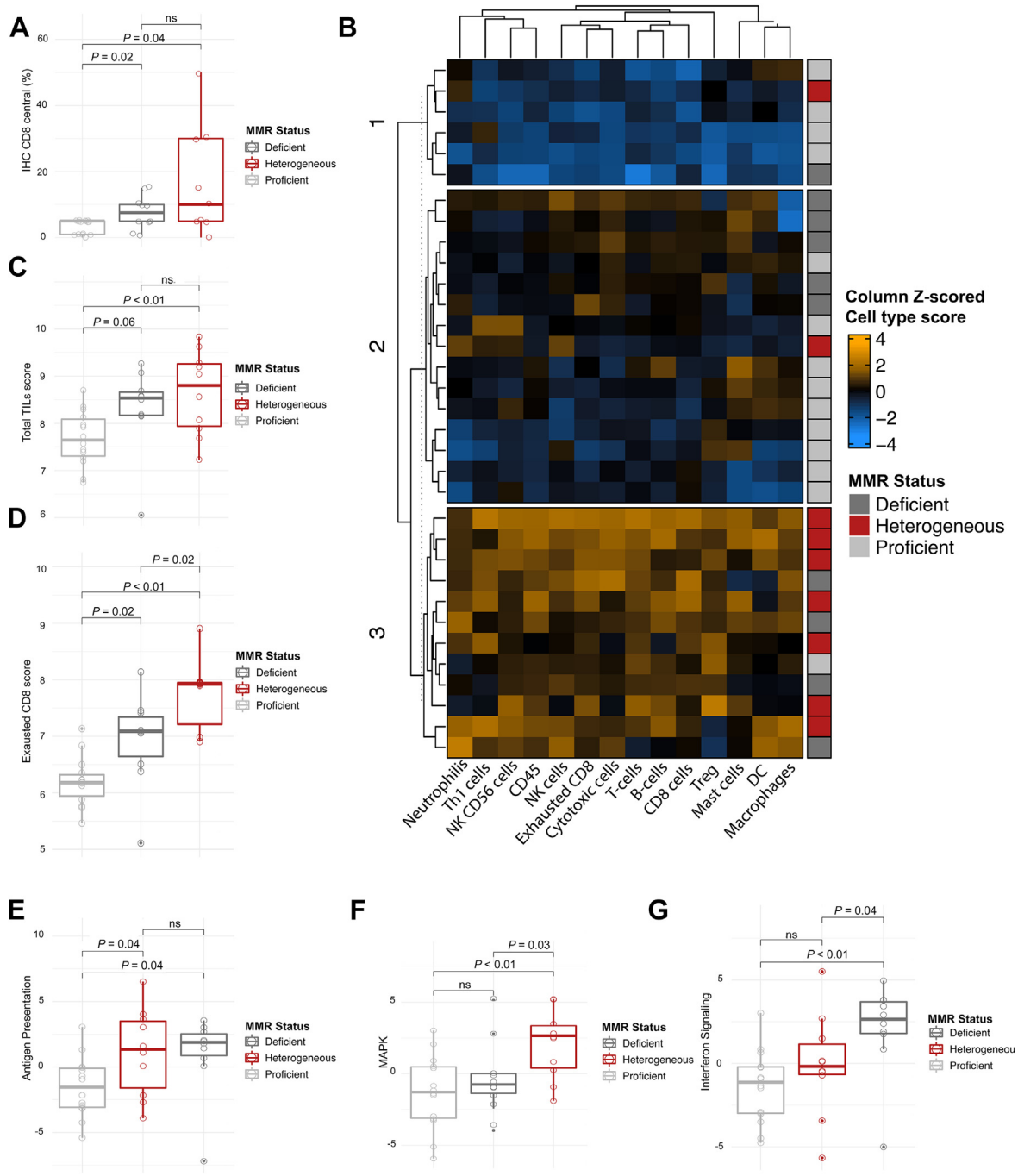

Figure 4.
Protein and gene expression features of the MMRh tumors. (A) A box plot of the $CD8^+$ TIL distribution across the MMRp, MMRd, and MMRh cohorts. (B) Unsupervised clustering of the immune cell type z scores for the analyzed tumors. The clustering identifies an MMRh-enriched group characterized by an enhancement of immune cells (number 3). (C) A box plot of the gene expression–based total amount of TILs across the MMRp, MMRd, and MMRh cohorts. (D) A box plot of gene expression–based scores for the total amount of exhausted CD8 cells across the MMRp, MMRd, and MMRh cohorts. (E) A box plot of gene expression–based antigen presentation pathway scores across the MMRp, MMRd, and MMRh cohorts. (F) A box plot of gene expression–based MAPK pathway scores across the MMRp, MMRd, and MMRh cohorts. (G) A box plot of gene expression–based interferon signaling pathway scores across the MMRp, MMRd, and MMRh cohorts. d, deficient; H, heterogeneous; IHC, immunohistochemistry; MMR, mismatch repair; ns, not significant ($P > .05$); p, proficient; TIL, tumor-infiltrating lymphocyte.

diagram showed a scenario similar to that of MMRh case 7, with the private variants in both the positive and negative components without significant differences in number (Supplementary Fig. S3B) but still characterized by subclonal VAFs ($P < .0001$, unpaired *t* test; Supplementary Fig. S2).

Taken together, our data show that in these MMRh double-loss cases, the microdissected areas with a double-negative IHC pattern are characterized by a higher level of non-synonymous variants and unstable microsatellite loci compared with the areas of the same tumor showing loss of expression of a

single MMR protein/heterodimer. In 4 tumors, the double-loss areas always showed the private variants, whereas in 3 cases, the independent variants were identified in both components. The private variants were subclonal compared with the shared ones.

Gene Expression Profiling of MMRh Tumors

To understand whether the increase in mutational load identified in the areas with different MMR protein expression would correspond to unique features of the tumor microenvironment, we performed both immunohistochemical analyses for TIL markers (CD8 and CD3) and a targeted gene expression analysis using a 770-gene panel (IO 360; NanoString Technologies) on the RNA extracted from the tumor bulk of all 9 MMRh tumors. The characterization of 15 MMRp and MSS CRCs and 10 MMRd and MSI CRCs allowed differential expression analysis.

The accurate evaluation of CD3/CD8 expression on TILs was performed in terms of both level of expression and topographic localization within the tumor. Using the percentage of positivity as a continuous value, the distribution analysis defined a significant difference for $CD8^+$ TILs, with an increased number of $CD8^+$ TILs in the core of the lesion (central tumor area) in MMRh and MMRd CRCs compared with MMRp lesions ($P = .04$ and $P = .01$, unpaired t test, respectively). The MMRh and MMRd CRCs were characterized by a similar median value of $CD8^+$ cells ($P = .32$, unpaired t test) but with a marked variance between the 2 groups ($P < .05$, Levene test; Fig. 4A).

We did not perform analysis of immune cell infiltration by separately counting TILs in the areas with distinct IHC patterns of MMRh cases because the stromal boundaries were not clear-cut and the analysis would have suffered from subjective assessment, thus precluding solid conclusions. Rather, we integrated our data using the gene expression profiling analysis, which allowed to increase the number of biomarkers, enabling a deeper stratification of the analyzed tumors. The first level of analysis concerned the characterization of the immunologic microenvironment. The immunophenotyping, obtained by calculating the immune cell type score, enabled to show that all MMRh tumors were enriched in immunologic markers and cells, thus forming a separate cluster with the exception of the MMRh case 6, the tumor characterized by the smallest ΔTMB within the heterogeneous cohort. Of note, the unsupervised clustering identified 3 groups of tumors with high infiltration, medium infiltration, and low infiltration (Fig. 4B). As expected, MMRp CRCs presented the lowest level of the cell type score. The MMRd lesions belonged to the medium-expression/high-expression group of immunophenotypic markers, with the MMRd case 1 showing an independent behavior.

The MMRh CRCs showed significantly higher levels of total TILs than the MMRp group ($P = .009$, unpaired t test). Of note, the TIL levels in MMRd tumors were also higher than those in MMRp tumors; however, the difference was borderline for significance ($P = .06$, unpaired t test). Considering all cell types, we identified an enrichment of the immune cell types in both MMRh and MMRd compared with MMRp. Of note, these enrichments were more pronounced for the MMRh, with 11 of 14 cell types increased compared with the MMRp, whereas 6 of 14 types were enriched in the MMRd than in MMRp (Supplementary Table S3 and Supplementary Fig. S4). Interestingly, the total amount of TILs and exhausted CD8 cells (but not total CD8 cells) was significantly higher in the MMRh than in both MMRp ($P < .01$, unpaired t test, for both cell types) and MMRd ($P = .047$ and $P = .023$, unpaired t test, respectively; Fig. 4C, D; Supplementary Table S3 and Supplementary Fig. S4).

Based on the pathway score, which summarized the expression level of each pathway based on predefined gene sets, MMRh and MMRd were enriched for immune-related pathways (myeloid and lymphoid compartments, cytokines and chemokines signaling, and the immune cell adhesin and migration; Supplementary Table S4 and Supplementary Fig. S5) and the antigen presentation pathway (Fig. 4E) compared with MMRp. An enhancement of the MAPK pathway was detected only in the MMRh ($P < .01$ and $P = .03$, unpaired t test, compared with MMRp and MMRd, respectively; Fig. 4F).

MMR-deficient lesions showed an increased interferon signaling ($P < .01$ and $P = .02$, unpaired t test, compared with MMRp and MMRH, respectively; Fig. 4G), and these tumors also displayed a lower expression of WNT and Hedgehog signaling than the proficient and heterogeneous ones (Supplementary Table S4 and Supplementary Fig. S5). Taken together, these data suggest that MMRh CRCs are independent tumor entities with an enhanced lymphocyte infiltration but characterized by the accumulation of exhausted $CD8^+$ cells and a reduced activation of the interferon signaling pathway.

Discussion

Molecular and cellular heterogeneity represents a key feature of tumors. In this study, we show a phenomenon of complex heterogeneity of MMR loss of expression in CRC, in which the spatial heterogeneity of 1 of the 4 main proteins (MLH1, MSH2, MSH6, or PMS2) was associated with loss of expression of another MMR dimer. This stemmed from the observation of "double-negative" areas at IHC assessment in some tumors that led to the identification of CRCs with double-loss of expression of MMR proteins in a subpopulation of tumor cells (MMRh double-loss CRCs). A next-generation sequencing analysis on the microdissected areas in these MMR heterogeneous cases demonstrated an accrual of unstable microsatellite sites in the tumor areas with double-loss of expression, accompanied by a significantly higher accumulation of subclonal somatic variants. These unique tumors showed a brisker lymphocytic infiltration than MSS CRCs and an increased expression of specific T-cell signature compared with the classical MMRd tumors. In addition, they harbored an increased level of exhausted $CD8^+$ T cells, thus suggesting an association of high hypermutability status and gradual deterioration of $CD8^+$ T cells.

The loss of the MMR system is encountered in approximately 15% in CRC, when considering both the germline Lynch syndrome and the sporadic setting, and identifies patients with genetic features of hypermutability and instability of microsatellite sequences. This trait can be identified by evaluating the phenotypic effect (ie, expression of the MLH1-PMS2 and MSH2-MSH6 heterodimers by IHC) or the molecular effect (ie, MSI assessed by DNA sequencing methods). The 2 tests may provide divergent results, in particular if tumor heterogeneity is present. The molecular test leading to MSI versus MSS status is routinely applied on the "bulk" tumor, thus reducing the spatial information.[21] The topographic observation of MMR protein expression by IHC allows to detect the presence and type of heterogeneity. There are rare scenarios with clear-cut features in which 2 distinct tumoral clones are present, one with retained MMR protein expression and MSS and the other with loss of MMR protein expression and MSI.[12,14,22,23]

More frequently (approximately 8%),[13] tumors can show a complex admixture of areas with retained and lost MMR expression, as recapitulated by the cases described in this study. In this

scenario, technical issues need to be verified as poor tissue fixation may lead to artifactual staining patterns,[24] and we made sure in our study that internal controls (positive stromal cells) were detectable in the tumor areas showing lack of MMR expression. Once technical issues are ruled out, true complexity featured by heterogeneous MMR loss of expression can be diagnosed. In some instances, heterogeneity of MMR expression of 1 heterodimer is associated with a complete loss of the other heterodimer, thus leading to tumors composed of areas lacking expression of all 4 MMR proteins. Jaffrelot et al[25] have recently identified 4 main groups of unusual conditions of MMR protein expression or MSI status. Of these, "group 4" was comprised mainly of sporadic CRC with areas harboring full loss of MMR protein expression. Previously, 4 cases with loss of MLH1 (because of promoter hypermethylation) and simultaneous loss of MSH6 have been reported, all characterized by the p.V600E variant of the *BRAF* gene.[16] These anecdotal cases have been reported but never studied in detail before, and, to our knowledge, our study represents the first attempt in this respect. We were able to microdissect 4 cases that pertained to the MMRh double-loss CRCs and show different molecular features pertaining to the distinct areas. The observation of differential hypermutability in spatially distinct areas demonstrates that some tumors harbor a sort of "super-MMRd" phenotype.

Our data suggest that a precise assessment of MMR protein expression in CRCs may be important, even in a scenario of MSI because of loss of 1 of the 2 heterodimers (MMRh double-loss CRCs), thus further stressing the importance of the IHC assessment of MMR machinery status. Overall, the complexity in MMR diagnostic assessment that we reveal may hold possible clinical implications with respect to prognostic stratification and prediction of response to therapy. Patients with MSI-H, early-stage CRC show better prognosis, with no clear benefit derived from adjuvant chemotherapy.[26] Patients with MSI-H, advanced-stage CRC are the optimal/ideal candidates to ICI, and this has exponentially increased its significance.[26] One may wonder whether a complex MSI scenario with "super-MMRd" phenotype may have an impact at both levels. In particular, the determinants of response to ICI are complex, and a fraction of patients with MSI CRC seems not to benefit from treatment.[10,11,27]

Interestingly, our gene expression data confirmed an accumulation of lymphocyte infiltrate[28] in MMRh and MMRd tumors. However, a significantly higher value of exhausted $CD8^+$ was revealed in MMRh tumors compared with MMRd in a context of superimposable total $CD8^+$ T-cell amount. These data are in line with the potential association between high level of DNA mutability of these tumors with a likely marked increase in neoantigenic load, which has been shown to induce $CD8^+$ T-cell exhaustion.[29] The presence of exhausted T cells has been reported as associated with reduced overall survival[30] and characterized by a "dualistic" nature in the context of the level of response to checkpoint inhibitors. On the one hand, the reversion of CD8 exhaustion has represented the basis of the original paradigm of checkpoint inhibition.[31] On the other hand, the definition of phenotypic heterogeneity of $CD8^+$ exhaustion[31] and the evidence regarding the reduced response to ICI of MSI CRCs rich in $CD8^+$ "terminal exhausted"[32] cells demand more in-depth mechanistic studies. We do not have direct evidence of response to treatment in our cohort. However, we could investigate in depth the pathway expression analysis, which demonstrated that both MMRh and MMRd tumors display an accumulation of genes associated with the antigen presentation, in line with MSI features. However, the interferon signaling pathway, recognized as the driver of spontaneous and ICI-induced antitumor immunity,[33,34] was significantly enriched in homogeneous MMRd rather than MMRh tumors. Along the same lines, a higher number of infiltrating cytotoxic T cells were observed in homogeneous MMRd rather than MMRh lesions. Cytotoxic T cells represent the most powerful effectors in the anticancer immune response and constitute the backbone of cancer immunotherapy.[35] Overall, our data may be hypothesis generating in postulating that these heterogeneous CRCs may be less responsive to ICIs because of the accumulation of exhausted $CD8^+$ cells over cytotoxic $CD8^+$ cells.

The present study has some limitations. We were not able to obtain sequencing data for all MMR heterogeneous tumors, mostly because of the impossibility of accurately separating by microdissection the differentially expressed areas.

One of the cases not suitable for microdissection was the only MMRh double-loss case that was MSI-low by polymerase chain reaction analysis and harbored a total of 11.1% of unstable loci by TruSight Oncology 500 analysis. Regrettably, we were not able to investigate in depth the molecular landscape of the distinct areas identified on the basis of the IHC staining.

Furthermore, the RNA analysis was performed on the "bulk" lesion and not at the single-cell level, preventing a more accurate classification of the immunophenotyping. Finally, we do not have correlative data with response to therapy and in particular to ICIs because this was a cohort of early CRC treated with surgery with or without chemotherapy. Analysis of metastatic CRC harboring these features would be of major interest to address as to whether this pattern may contribute a differential response to ICIs. Regrettably, we did not have access to cohorts of patients with metastatic CRC with available tissue samples and treated with ICIs to assess the possible correlations.

Overall, the study characterizes a novel tumor heterogeneity model in CRC, with variable degrees of MMR protein expression. Tumors featuring areas with MMR "double-negativity" at the IHC test leading to full loss of MMR proteins ("MMRh double-loss CRCs") were associated with an accumulation of somatic variants in the double-loss areas and showed a lymphocytic infiltrate enriched for exhausted T cells, identifying a subset of "super-MMRd" CRCs.

Acknowledgments

This article is dedicated to the memory of Dr Laura Annaratone.

Author Contributions

C.M., E.Berrino, and E.Bonoldi provided the study concept. E.Berrino, M.C.A., M.G., and E.V. curated the data. E.Berrino and C.M. conducted a formal analysis. A.B. and C.M. acquired funding. E.Berrino, C.M., E.Bonoldi, M.C.A., E.F., S.M., V.A., G.G., A.B., S.M., and K.G. conducted investigation. C.M., A.S., and A.B. provided the resources. E.Berrino provided the software. C.M., E.Bonoldi, S.M., A.S.-B., S.S., and A.B. supervised the study. E.Berrino conducted validation. E.Berrino conducted visualization. E.Berrino, C.M., and E.Bonoldi drafted the original manuscript. All authors reviewed and edited the manuscript.

Data Availability

All data are included in the main article and the supplementary material.

Funding

This research was funded by FONDAZIONE AIRC (Associazione Italiana per la Ricerca sul Cancro) under 5 per Mille 2018—ID 21091 program—PI: B.A. and GL: M.C. E.Berrino was the recipient of a PhD fellowship under the funding of Dipartimenti di Eccellenza 2018-2022 (project number: 521 D15D18000410001). A.B. and E.F. were the recipients of FPRC (Fondazione Piemontese per la Ricerca sul Cancro) 5×1000 Ministero Salute 2016, Progetto Ardite. Investigators at Grande Ospedale Metropolitano Niguarda and Università degli Studi di Milano were also funded by Fondazione Oncologia Niguarda. G.G. is supported by FPRC 5×mille 2017 Ministero Salute PTCRC-Intra 2020 (REGENERATION-YIG 2020 project) and AIRC MFAG 2019 grant ID 24604—PI: G.G.

Declaration of Competing Interest

C.M. reports personal consultancy fees from Bayer, Roche, AstraZeneca, and Daiichi Sankyo. A.B. served in a consulting/advisory role for Illumina and Inivata. A.B. and G.G. are cofounders and shareholders of NeoPhore. A.B. is a member of the NeoPhore Scientific Advisory Board. E.F. reports personal consultancy fees from Merck. A.S.-B. is a member of advisory boards for Amgen, Bayer, Sanofi, and Servier. S.S. is an advisory board member for AstraZeneca, Bayer, BMS, CheckmAb, Daiichi Sankyo, Menarini, Merck, and Seattle Genetics. All remaining authors have no conflicts of interest to declare.

Ethics Approval and Consent to Participate

Patients included in the study signed the informed consent within the AlfaOmega project (approved by the Ethical Committee of the Niguarda Cancer Center on April 7, 2020, #145-07042020; NCT04120935 and NCT05101382).

Supplementary Material

The online version contains supplementary material available at https://doi.org/10.1016/j.modpat.2022.100012.

References

1. Germano G, Amirouchene-Angelozzi N, Rospo G, Bardelli A. The clinical impact of the genomic landscape of mismatch repair-deficient cancers. *Cancer Discov*. 2018;8(12):1518–1528.
2. Jiricny J. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol*. 2006;7(5):335–346.
3. Li GM. Mechanisms and functions of DNA mismatch repair. *Cell Res*. 2008;18(1):85–98.
4. Lynch HT, Snyder CL, Shaw TG, Heinen CD, Hitchins MP. Milestones of Lynch syndrome: 1895-2015. *Nat Rev Cancer*. 2015;15(3):181–194.
5. Tubbs A, Nussenzweig A. Endogenous DNA damage as a source of genomic instability in cancer. *Cell*. 2017;168(4):644–656.
6. Turajlic S, Litchfield K, Xu H, et al. Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. *Lancet Oncol*. 2017;18(8):1009–1021.
7. Hause RJ, Pritchard CC, Shendure J, Salipante SJ. Classification and characterization of microsatellite instability across 18 cancer types. *Nat Med*. 2016;22(11):1342–1350.
8. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. *J Clin Oncol*. 2005;23(3):609–618.
9. Amodio V, Mauri G, Reilly NM, et al. Mechanisms of immune escape and resistance to checkpoint inhibitor therapies in mismatch repair deficient metastatic colorectal cancers. *Cancers (Basel)*. 2021;13(11):2638.
10. Le DT, Durham JN, Smith KN, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science*. 2017;357(6349):409–413.
11. Le DT, Uram JN, Wang H, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med*. 2015;372(26):2509–2520.
12. Joost P, Veurink N, Holck S, et al. Heterogenous mismatch-repair status in colorectal cancer. *Diagn Pathol*. 2014;9:126.
13. Watson N, Grieu F, Morris M, et al. Heterogeneous staining for mismatch repair proteins during population-based prescreening for hereditary non-polyposis colorectal cancer. *J Mol Diagn*. 2007;9(4):472–478.
14. Loupakis F, Maddalena G, Depetris I, et al. Treatment with checkpoint inhibitors in a metastatic colorectal cancer patient with molecular and immunohistochemical heterogeneity in MSI/dMMR status. *J Immunother Cancer*. 2019;7(1):297.
15. Pai RK, Plesec TP, Abdul Karim FW, et al. Abrupt loss of MLH1 and PMS2 expression in endometrial carcinoma: molecular and morphologic analysis of 6 cases. *Am J Surg Pathol*. 2015;39(7):993–999.
16. Westwood A, Glover A, Hutchins G, et al. Additional loss of MSH2 and MSH6 expression in sporadic deficient mismatch repair colorectal cancer due to MLH1 promoter hypermethylation. *J Clin Pathol*. 2019;72(6):443–447.
17. Berrino E, Filippi R, Visintin C, et al. Collision of germline POLE and PMS2 variants in a young patient treated with immune checkpoint inhibitors. *NPJ Precis Oncol*. 2022;6(1):15.
18. Berrino E, Annaratone L, Bellomo SE, et al. Integrative genomic and transcriptomic analyses illuminate the ontology of HER2-low breast carcinomas. *Genome Med*. 2022;14(1):98.
19. Zhang H, Wang Y, Ji Q, et al. Clinicopathological and molecular characteristics of patients with hypermutant lung cancer: a retrospective cohort study. *Oncol Lett*. 2021;21(4):329.
20. Campbell BB, Light N, Fabrizio D, et al. Comprehensive analysis of hypermutation in human cancer. *Cell*. 2017;171(5):1042–1056.e10.
21. Kim JH, Kang GH. Molecular and prognostic heterogeneity of microsatellite-unstable colorectal cancer. *World J Gastroenterol*. 2014;20(15):4230–4243.
22. Loupakis F, Depetris I, Biason P, et al. Prediction of benefit from checkpoint inhibitors in mismatch repair deficient metastatic colorectal cancer: role of tumor infiltrating lymphocytes. *Oncologist*. 2020;25(6):481–487.
23. McCarthy AJ, Capo-Chichi JM, Spence T, et al. Heterogenous loss of mismatch repair (MMR) protein expression: a challenge for immunohistochemical interpretation and microsatellite instability (MSI) evaluation. *J Pathol Clin Res*. 2019;5(2):115–129.
24. Greenberg A, Kariv R, Solar I, Hershkovitz D. Geographic heterogeneity for mismatch repair proteins is associated with defects in DNA repair. *Isr Med Assoc J*. 2020;22(1):32–36.
25. Jaffrelot M, Farés N, Brunac AC, et al. An unusual phenotype occurs in 15% of mismatch repair-deficient tumors and is associated with non-colorectal cancers and genetic syndromes. *Mod Pathol*. 2022;35(3):427–437.
26. Schiappacasse Cocio GV, Schiappacasse ED. Is adjuvant chemotherapy efficient in colon cancer with high microsatellite instability? A look towards the future. *Cancer Res*. 2019;79(3):441–444.
27. Overman MJ, McDermott R, Leach JL, et al. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *Lancet Oncol*. 2017;18(9):1182–1191.
28. Smyrk TC, Watson P, Kaul K, Lynch HT. Tumor-infiltrating lymphocytes are a marker for microsatellite instability in colorectal carcinoma. *Cancer*. 2001;91(12):2417–2422.
29. Lopez de Rodas M, Schalper KA. Tumour antigen-induced T cell exhaustion—the archenemy of immune-hot malignancies. *Nat Rev Clin Oncol*. 2021;18(12):749–750.
30. Sorrentino C, D'Antonio L, Fieni C, Ciummo SL, Di Carlo E. Colorectal cancer-associated immune exhaustion involves T and B lymphocytes and conventional NK cells and correlates with a shorter overall survival. *Front Immunol*. 2021;12(2):778329.
31. Budimir N, Thomas GD, Dolina JS, Salek-Ardakani S. Reversing T-cell exhaustion in cancer: lessons learned from PD-1/PD-L1 immune checkpoint blockade. *Cancer Immunol Res*. 2022;10(2):146–153.
32. Tian S, Wang F, Zhang R, Chen G. Global pattern of $CD8^+$ T-cell infiltration and exhaustion in colorectal cancer predicts cancer immunotherapy response. *Front Pharmacol*. 2021;12:715721.
33. Curtsinger JM, Agarwal P, Lins DC, Mescher MF. Autocrine IFN-gamma promotes naive CD8 T cell differentiation and synergizes with IFN-alpha to stimulate strong function. *J Immunol*. 2012;189(2):659–668.
34. Maraskovsky E, Chen WF, Shortman K. IL-2 and IFN-gamma are two necessary lymphokines in the development of cytolytic T cells. *J Immunol*. 1989;143(4):1210–1214.
35. Raskov H, Orhan A, Christensen JP, Gogenur I. Cytotoxic $CD8^+$ T cells in cancer and cancer immunotherapy. *Br J Cancer*. 2021;124(2):359–367.