Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Le DT, Uram JN, Wang H, et al. PD-1 blockade in tumors with mismatch-repair deficiency. N Engl J Med 2015;372:2509-20. DOI: 10.1056/NEJMoa1500596

(Table S3 in Excel file updated March 1, 2018)

Amgen Exhibit 1032
Amgen v. BMS

Supplementary Appendix

Table of Contents to the Supplementary Appendix

I. LIST OF INVESTIGATORS 2
II. SUPPLEMENTARY METHODS 3
III. SUPPLEMENTARY FIGURES 9
Figure S1. Overview of Study Design 10
Figure S2. Spider plot of radiographic response 11
Figure S3. MMR-proficient and deficient CRCs have comparable time on treatment and duration of metastatic disease prior to study enrollment 12
Figure S4. Waterfall plot of biochemical response 13
Figure S5. Somatic mutations in MMR-deficient and proficient tumors. 14
Figure S6. Immunohistochemistry of CD8 and PD-L1 Expression 15
Figure S7. CD8 and PD-L1 Expression in the MMR-deficient and MMR-proficient tumor microenvironment 16
Figure S8. CD8 expression and clinical benefit to pembrolizumab 17
IV. SUPPLEMENTARY TABLES 18
Table S1. Comparison of immune-related and RECIST response criteria 19
Table S2. Immune-Related response to treatment 20
Table S3. Summary of genomic and IHC data 21
Table S4. Correlation of total somatic mutations and mutation associated neoantigens (MANA) with clinical outcomes 22
Table S5. Correlation of immune markers with clinical outcome 23
V. REFERENCES 24

I. LIST OF INVESTIGATORS

Dung T. Le
Jennifer N. Uram
Hao Wang
Bjarne R. Bartlett
Holly Kemberling
Aleksandra D. Eyring
Andrew D. Skora
Brandon S. Luber
Nilofer S. Azad
Dan Laheru
Barbara Biedrzycki
Ross C. Donehower
Atif Zaheer
George A. Fisher
Todd S. Crocenzi
James J. Lee
Steven M. Duffy
Richard M. Goldberg
Albert de la Chapelle
Minori R. Koshiji
Feriyl Bhaijee
Thomas Huebner
Ralph H. Hruban
Laura D. Wood
Nathan Cuka
Drew Pardoll
Nickolas Papadopoulos
Kenneth W. Kinzler
Shibin Zhou
Toby C. Cornish
Janis M. Taube
Robert A. Anders
James R. Eshleman
Bert Vogelstein
Luis A. Diaz, Jr

This material may be protected by Copyright law (Title 17 U.S. Code)

II. SUPPLEMENTARY METHODS

PATIENTS

To be eligible for participation in this study, patients had to be at least 18 years of age, have histologically confirmed evidence of previously-treated, progressive carcinoma. All patients underwent MMR status testing prior to enrollment. All patients had at least one measurable lesion as defined by the Response Evaluation Criteria in Solid Tumors (RECIST), version 1.1, an Eastern Cooperative Oncology Group (ECOG) performance-status score of 0 or 1, and adequate hematologic, hepatic, and renal function. Eligible patients with CRC must have received at least 2 prior cancer therapies and patients with other cancer types must have received at least 1 prior cancer therapy. Patients with untreated brain metastases, history of HIV, hepatitis B, hepatitis C, clinically significant ascites/effusions, or autoimmune disease were excluded.

STUDY OVERSIGHT

Initial drafts of the manuscript were prepared by a subset of the authors and all authors contributed to the final manuscript. All the authors made the decision to submit the manuscript for publication. The principal investigator and study sponsor vouch for the accuracy and completeness of the data reported as well as adherence to the protocol.

HLA TYPING

HLA-A, HLA-B and HLA-C Sequence Based Typing can be divided into three distinct steps, as described below. A generic, A*02 specific, B generic, B group specific, C generic and C*07 specific PCR and sequencing mixes were made in the JHU core facility. Celera's AlleleSEQR HLA-B Sequence Based Typing kit was used for B generic SBT. The HLA-A typing scheme is composed of two PCR reactions, A generic and A*02 specific. A generic amplicon encompasses partial exon 1- partial exon 5. A*02 amplicon encompasses partial intron 1 - partial exon 5. HLA-B typing scheme is composed of two PCR reactions, B generic and B group specific. The B generic PCR is a multiplexed reaction containing two PCR amplicons encompassing exon 2 – exon 3 and exon 4 – exon 7. B group specific amplicon encompasses partial intron 1 - partial exon 5. HLA-C typing scheme is composed of two PCR reactions, C generic and C*07 specific. C generic and C*07 specific amplicons encompasses exons 1 - 7.

The specificity of the HLA-A and B PCR employed AmpliTaq Gold DNA polymerase. The GeneAmp High Fidelity enzyme is used for the HLA-C and C*07 PCR mixes. This enzyme is a mix of two polymerases: AmpliTaq DNA polymerase (non-proofreading polymerase) and a proofreading polymerase. This enzyme mix is necessary to produce efficient and robust amplification of the larger full length HLA-C amplicon.

PCR product purification was performed using Exonuclease I and Shrimp Alkaline Phosphatase The A generic and B generic amplicons were bi-directionally sequenced for exons 2,3,4. The C generic amplicon was bi-directionally sequenced for exons 2,3 and sequenced in a single direction for exons 1,4,5,6,7. A*02 specific, B group specific and C*07 specific amplicons were sequenced in a single direction for exons 2,3. All sequencing reactions were performed with Big Dye Terminator V1.1 from Applied Biosystems and sequenced with an ABI Prism 3500XL Genetic Analyzer. Conexio Genomic's "Assign SBT" allele assignment software was used to process the data files.

MISMATCH REPAIR STATUS TESTING[1,2]

Six slides of tumor and normal (uninvolved lymph node or margin of resection) were cut (5 microns each), deparaffinized (xylene), and one stained with hematoxylin and eosin (H+E). A tumor area containing at least 20% neoplastic cells, designated by a board-certified Anatomic Pathologist was macrodissected using the Pinpoint DNA isolation system (Zymo Research, Irvine, CA), digested in proteinase K for 8 hours and DNA was isolated using a QIAamp DNA Mini Kit (Qiagen, Valencia, CA). MSI was assessed using the MSI Analysis System (Promega, Madison, WI), composed of 5 pseudomonomorphic mononucleotide repeats (BAT-25, BAT-26, NR-21, NR-24 and MONO-27) to detect MSI and 2-pentanucleotide repeat loci (PentaC and PentaD) to confirm identity between normal and tumor samples, per manufacturer's instructions. Following amplification of 50-100 ng DNA, the fluorescent PCR products were sized on an Applied Biosystems 3130xl capillary electrophoresis instrument (Invitrogen, Calsbad, CA). Pentanucleotide loci confirmed identity in all cases. Controls included water as a negative control and a mixture of 80% germline DNA with 20% MSI cancer DNA as a positive control. The size in bases was determined for each microsatellite locus and tumors were designated as MSI if two or more mononucleotide loci varied in length compared to the germline DNA.

SEQUENCING ANALYSIS

Samples

Samples provided as FFPE blocks or frozen tissue underwent pathological review to determine tumor cellularity. Tumors were macrodissected to remove contaminating normal tissue, resulting in samples containing >20% neoplastic cells. Matched normal samples were provided as blood, saliva or normal tissue obtained from surgery.

Sample Preparation and Next-Generation Sequencing[3]

Sample preparation, library construction, exome capture, next generation sequencing, and bioinformatics analyses of tumor and normal samples were performed at Personal Genome Diagnostics, Inc. (Baltimore, Maryland). In brief, DNA was extracted from frozen or formalin-fixed paraffin embedded (FFPE) tissue, along with matched blood or saliva samples using the Qiagen DNA FFPE tissue kit or Qiagen DNA blood mini kit (Qiagen, CA). Genomic DNA from tumor and normal samples were fragmented and used for Illumina TruSeq library construction (Illumina, San Diego, CA) according to the manufacturer's instructions or as previously

described[4]. Briefly, 50 nanograms (ng) - 3 micrograms (µg) of genomic DNA in 100 microliters (µl) of TE was fragmented in a Covaris sonicator (Covaris, Woburn, MA) to a size of 150-450bp. To remove fragments smaller than 150bp, DNA was purified using Agencourt AMPure XP beads (Beckman Coulter, IN) in a ratio of 1.0 to 0.9 of PCR product to beads twice and washed using 70% ethanol per the manufacturer's instructions. Purified, fragmented DNA was mixed with 36 µl of H2O, 10 µl of End Repair Reaction Buffer, 5 µl of End Repair Enzyme Mix (cat# E6050, NEB, Ipswich, MA). The 100 µl end-repair mixture was incubated at 20°C for 30 min, and purified using Agencourt AMPure XP beads (Beckman Coulter, IN) in a ratio of 1.0 to 1.25 of PCR product to beads and washed using 70% ethanol per the manufacturer's instructions. To A-tail, 42 µl of end-repaired DNA was mixed with 5 µl of 10X dA Tailing Reaction Buffer and 3 µl of Klenow (exo-)(cat# E6053, NEB, Ipswich, MA). The 50 µl mixture was incubated at 37°C for 30 min and purified using Agencourt AMPure XP beads (Beckman Coulter, IN) in a ratio of 1.0 to 1.0 of PCR product to beads and washed using 70% ethanol per the manufacturer's instructions. For adaptor ligation, 25 µl of A-tailed DNA was mixed with 6.7 µl of H2O, 3.3 µl of PE-adaptor (Illumina), 10 µl of 5X Ligation buffer and 5 µl of Quick T4 DNA ligase (cat# E6056, NEB, Ipswich, MA). The ligation mixture was incubated at 20°C for 15 min and purified using Agencourt AMPure XP beads (Beckman Coulter, IN) in a ratio of 1.0 to 0.95 and 1.0 of PCR product to beads twice and washed using 70% ethanol per the manufacturer's instructions. To obtain an amplified library, twelve PCRs of 25 µl each were set up, each including 15.5 µl of H2O, 5 µl of 5 x Phusion HF buffer, 0.5 µl of a dNTP mix containing 10 mM of each dNTP, 1.25 µl of DMSO, 0.25 µl of Illumina PE primer #1, 0.25 µl of Illumina PE primer #2, 0.25 µl of Hotstart Phusion polymerase, and 2 µl of the DNA. The PCR program used was: 98°C for 2 minutes; 12 cycles of 98°C for 15 seconds, 65°C for 30 seconds, 72°C for 30 seconds; and 72°C for 5 min. DNA was purified using Agencourt AMPure XP beads (Beckman Coulter, IN) in a ratio of 1.0 to 1.0 of PCR product to beads and washed using 70% ethanol per the manufacturer's instructions. Exonic or targeted regions were captured in solution using the Agilent SureSelect v.4 kit according to the manufacturer's instructions (Agilent, Santa Clara, CA). The captured library was then purified with a Qiagen MinElute column purification kit and eluted in 17 µl of 70°C EB to obtain 15 µl of captured DNA library. (5) The captured DNA library was amplified in the following way: Eight 30uL PCR reactions each containing 19 µl of H2O, 6 µl of 5 x Phusion HF buffer, 0.6 µl of 10 mM dNTP, 1.5 µl of DMSO, 0.30 µl of Illumina PE primer #1, 0.30µl of Illumina PE primer #2, 0.30 µl of Hotstart Phusion polymerase, and 2 µl of captured exome library were set up. The PCR program used was: 98°C for 30 seconds; 14 cycles (exome) or 16 cycles (targeted) of 98°C for 10 seconds, 65°C for 30 seconds, 72°C for 30 seconds; and 72°C for 5 min. To purify PCR products, a NucleoSpin Extract II purification kit (Macherey-Nagel, PA) was used following the manufacturer's instructions. Paired-end sequencing, resulting in 100 bases from each end of the fragments for exome libraries and 150 bases from each end of the fragment for targeted libraries, was performed using Illumina HiSeq 2000/2500 and Illumina MiSeq instrumentation (Illumina, San Diego, CA).

Primary Processing of Next-Generation Sequencing Data and Identification of Putative Somatic Mutations[3]

Somatic mutations were identified using VariantDx custom software (Personal Genome Diagnostics, Baltimore, Maryland) for identifying mutations in matched tumor and normal

samples. Prior to mutation calling, primary processing of sequence data for both tumor and normal samples were performed using Illumina CASAVA software (v1.8), including masking of adapter sequences. Sequence reads were aligned against the human reference genome (version hg18) using ELAND with additional realignment of select regions using the Needleman-Wunsch method [5]. Candidate somatic mutations, consisting of point mutations, insertions, and deletions were then identified using VariantDx across the either the whole exome or regions of interest. VariantDx examines sequence alignments of tumor samples against a matched normal while applying filters to exclude alignment and sequencing artifacts. In brief, an alignment filter was applied to exclude quality failed reads, unpaired reads, and poorly mapped reads in the tumor. A base quality filter was applied to limit inclusion of bases with reported phred quality score > 30 for the tumor and > 20 for the normal. A mutation in the tumor was identified as a candidate somatic mutation only when (i) distinct paired reads contained the mutation in the tumor; (ii) the number of distinct paired reads containing a particular mutation in the tumor was at least 10% of read pairs; (iii) the mismatched base was not present in >1% of the reads in the matched normal sample as well as not present in a custom database of common germline variants derived from dbSNP; and (iv) the position was covered in both the tumor and normal at > 150X. Mutations arising from misplaced genome alignments, including paralogous sequences, were identified and excluded by searching the reference genome.

Candidate somatic mutations were further filtered based on gene annotation to identify those occurring in protein-coding regions. Functional consequences were predicted using snpEff and a custom database of CCDS, RefSeq and Ensembl annotations using the latest transcript versions available on hg18 from UCSC (https://genome.ucsc.edu/). Predictions were ordered to prefer transcripts with canonical start and stop codons and CCDS or Refseq transcripts over Ensembl when available. Finally mutations were filtered to exclude intronic and silent changes, while retaining mutations resulting in missense mutations, nonsense mutations, frameshifts, or splice site alterations. A manual visual inspection step was used to further remove artifactual changes.

MUTANT PEPTIDE MHC BINDING PREDICTION

Somatic frameshift, insertions, deletions, and missense mutations predicted to result in an amino acid change were analyzed for potential MHC class I binding based on the individual patient's HLA haplotype. Our initial analysis focused on HLA-A and HLA-B. Amino acid mutations were linked to their corresponding CCDS accession number and in instances where this was unavailable, either a Refseq or ensemble transcript was used to extract the protein sequence. To identify 8mer, 9mer, and 10mer epitopes, amino acid fragments surrounding each mutation were identified. These 15, 17, and 19 mutant amino acid fragments were analyzed by the epitope prediction program NetMHC 3.4.[6] Epitopes with a predicted affinity of <50nm were considered to be strong potential binders and epitopes with a predicted affinity of <500nm were considered to be weak potential binders as suggested by the NetMHC group[6].

To further refine the total neoantigen burden, we repeated that same process for the complementary wild-type peptide for each mutant peptide. We then filtered for mutant peptides that were strong potential binders when the complementary wild-type peptide was predicted a

weak potential binder. These mutant peptides are referred to as mutation-associated neoantigens (MANA). In the event that a patient had a (e.g., cases 1, 17 and 21) single MHC haplotype not supported by NetMHC 3.4, the individual haplotype was not included in our analysis.

STATISTICAL METHODS

Design of the trial[7]

This trial was conducted using a parallel two-stage design to simultaneously evaluate the efficacy of MK-3475 and MSI as a treatment selection marker for anti-PD-1 therapy. It consisted of two-stage phase 2 studies in parallel in the three cohorts of patients described in the text. The study agent, MK-3475, was administered at 10 mg/kg intravenously every 14 days.

For each of Cohort A and B, the co-primary endpoints were progression-free-survival (irPFS) at 20 weeks and objective response (irOR) assessed using immune related criteria. A step-down gatekeeping procedure was used to preserve the overall type I error. A two-stage Green-Dahlberg design was used to evaluate irPFS, with interim and final analysis after 15 and 25 patients, respectively. At stage 1, ≥ 1 of 15 free-of-progression at 20 weeks were required to proceed to the second stage, and ≥ 4 of 25 free-of-progression at 20 weeks were then required to proceed to test for irOR, with ≥ 4 of 25 responders (irCR or irPR) indicating promising efficacy in that cohort. Each cohort could be terminated for efficacy as soon as ≥ 4 free-of-progression at 20 weeks and ≥ 4 responses were confirmed, or be terminated for futility as soon as 0 of 15 in stage 1 were free-of-progression at 20 weeks or ≥ 22 subjects had disease progression by 20 weeks. This design achieves 90% power to detect a 20-week irPFS rate of 25% and 80% power to detect an irOR rate (irORR) of 21%, with an overall type I error of 0.05 at the null hypothesis of 20-week irPFS rate of 5% and irORR of 5%.

For Cohort C, the primary endpoint was irPFS at 20 weeks. A two-stage Green-Dahlberg two-stage design was used, with an interim and final analysis after 14 and 21 patients; at stage 1, ≥ 1 of 14 free-of-progression at 20 weeks were required to proceed to the second stage, with ≥ 4 of 21 free-of-progression at 20 weeks at the end indicating adequate efficacy in Cohort C. The cohort could be terminated as soon as ≥ 4 free-of-progression at 20 weeks were confirmed. The design has 81% power to detect a 20-week irPFS rate of 25% with a 5% type I error at the null hypothesis of 20-week irPFS rate of 5%.

Statistical analysis

Response and progression were evaluated using RECIST v1.1 and the immune-related response criteria (irRC) adopted from Wolchok et al.[8], which uses the sum of the products of bidimensional tumor measurements and incorporates new lesions into the sum. Progression-free survival (PFS) rates and irPFS rate at 20-weeks was estimated as the proportion of patients who were free-of-disease progression and alive at 20 weeks after the initiation of pembrolizumab. Patients who had disease progression prior to 20 weeks or were enrolled for >20 weeks at the time the study data were collated were included in the analysis for estimating 20-week PFS (irPFS) rate. Patients who dropped out early due to toxicities or worsening

disease and therefore did not have 20-week tumor assessment were considered as having progressive disease. ORR (irORR) was the proportion of patients who achieved best overall response of CR or PR (irCR or irPR). Patients who were in the study long enough to have tumor response evaluations were included in the analysis for estimating response rates. Among those who responded (CR or PR), duration of response was the time of first RECIST response to the time of disease progression, and was censored at the last evaluable tumor assessment for responders who had not progressed.

PFS and irPFS were defined as the time from the date of initial dose to the date of disease progression or the date of death due to any cause, whichever occurred first. PFS and irPFS were censored on the date of the last evaluable tumor assessment documenting absence of progressive disease for patients who were alive and progression-free. Overall survival (OS) was defined as the time from the date of initial dose to death due to any cause. For patients who were still alive at the time of analysis, the OS time was censored on the last date the patients were known to be alive. Survival times were summarized by the Kaplan-Meier method. As a *post hoc* analysis, log-rank tests were used to compare Cohort A and B and hazard ratios were estimated based on Cox models.

The association of percent CEA decline after 1 cycle with PFS or OS was assessed using landmark analysis based on Cox regression models. For correlative studies, non-parametric Wilcoxon test was used to compare mutational load between MMR-deficient and MMR-proficient patients. The effects of baseline mutational burden and immune markers on response and survival times were examined using logistic regression and Cox regression, respectively.

IMMUNOHISTOCHEMISTRY & IMAGE ANALYSIS

The fraction of malignant cells exhibiting a membranous pattern of B7-H1 expression and the percentage at the invasive front were quantified by three pathologists (R.A.A., F.B., and J.M.T.) as previously reported[9,10]. Image analysis was used to determine the number of CD8 diaminobenzidine (DAB)-stained cells. Using the H&E-stained slide for each case, we identified the following regions: i) tumor, ii) invasive front (the boundary between malignant and non-malignant tissue), and iii) normal tissue. The CD8-stained slides were scanned at 20x equivalent magnification (0.49 micrometers per pixel) on an Aperio ScanScope AT. Regions corresponding to tumor, invasive front and normal tissue (above, from the H&E) were annotated on separate layers using Aperio ImageScope v12.1.0.5029.

CD8-positive lymphocyte density was calculated in each of the above regions using a custom algorithm implemented in PIP[11]. Results were converted to Deepzoom images using the VIPS library[12] and visualized using the OpenSeadragon viewer (http://openseadragon.github.io).

III. SUPPLEMENTARY FIGURES

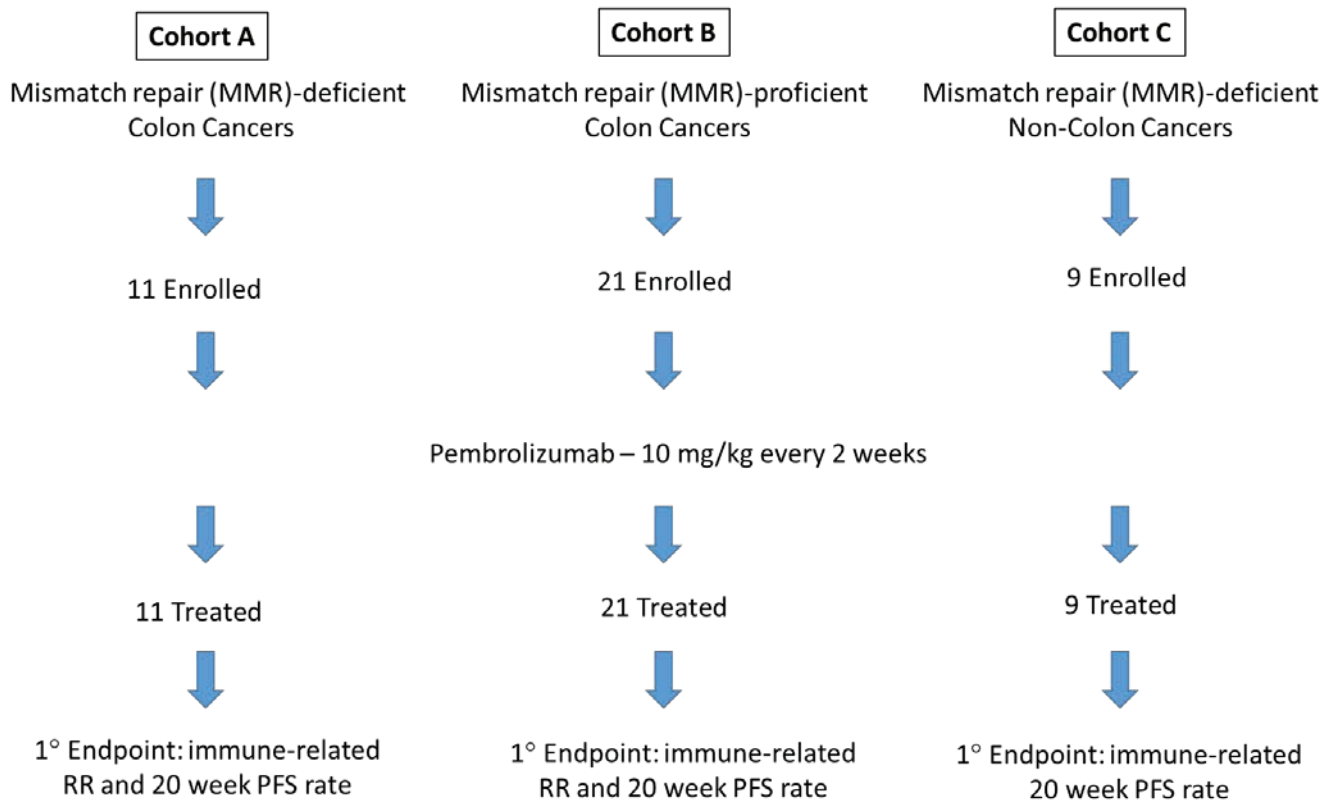


Figure S1. Overview of Study Design. MMR-deficiency, immune-related progression-free survival (PFS), response rate (RR). The patients who were in the study long enough to have tumor response evaluations were included in the analysis for estimating irORR (n=10, 18, and 7 in Cohort A, B, and C, respectively). The patients who were enrolled for >20 weeks at the time the study data were collated were included in the analysis for estimating 20 week irPFS rate (n=9, 18, and 6 in Cohort A, B, and C, respectively). All treated patients were included in the survival analysis of PFS and OS.

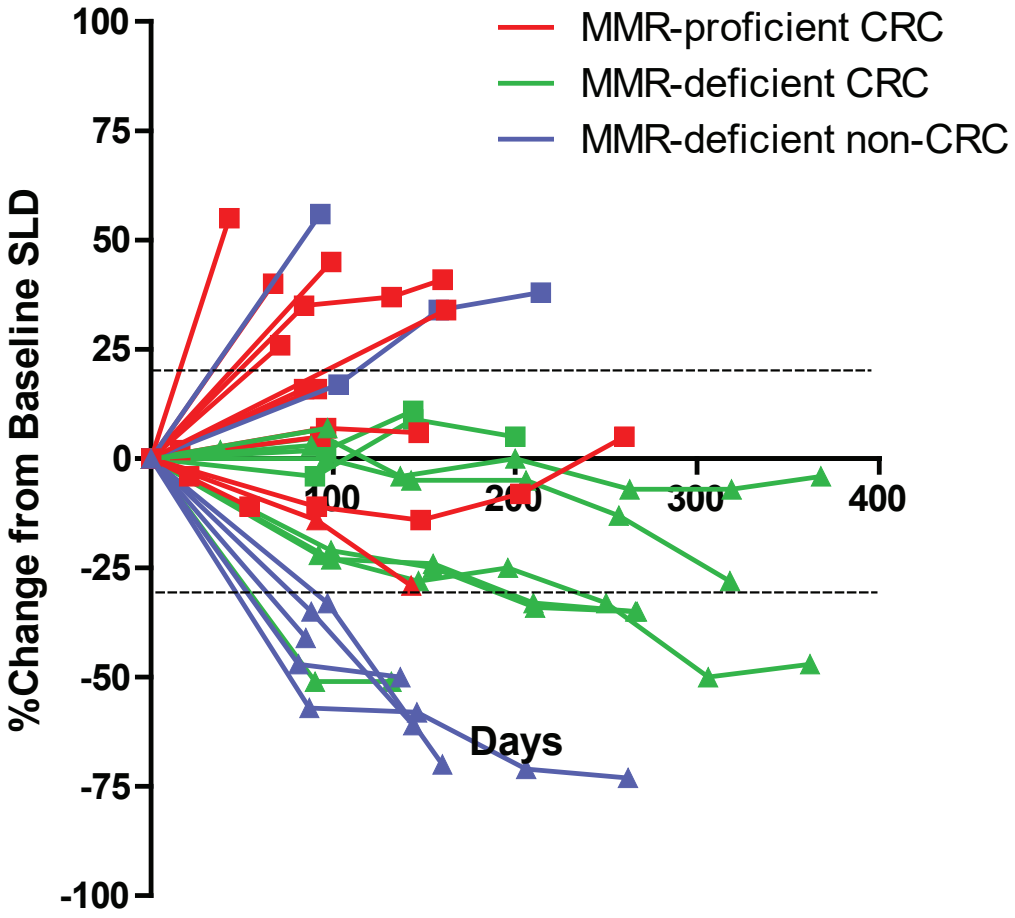
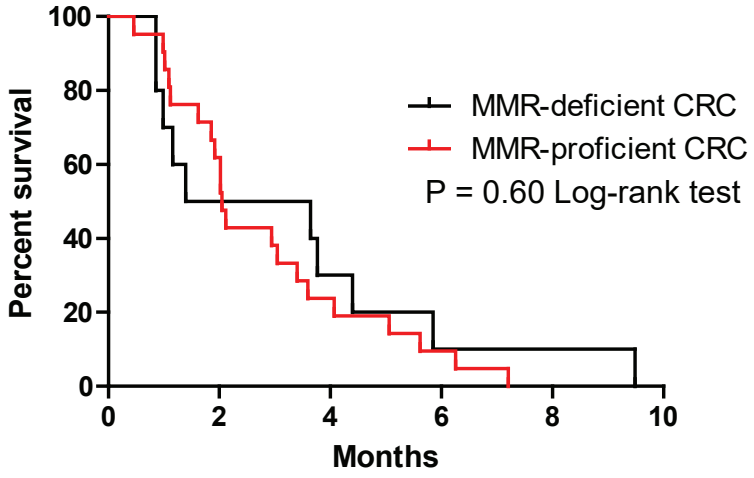


Figure S2. Spider plot of radiographic response. Tumor responses were measured at regular intervals and values show percent change of the sum of longest diameters (SLD) from the baseline measurements of each measurable tumor. Patients were only included if baseline and on study treatment scans were available. Green and red represent patients with MMR-deficient and proficient CRCs, respectively. Blue represents patients with MMR-deficient cancers other than CRC.

A. Time on therapy prior to enrollment on current study



B. Duration of the diagnosis of metastatic diagnosis prior to enrollment of pembrolizumab

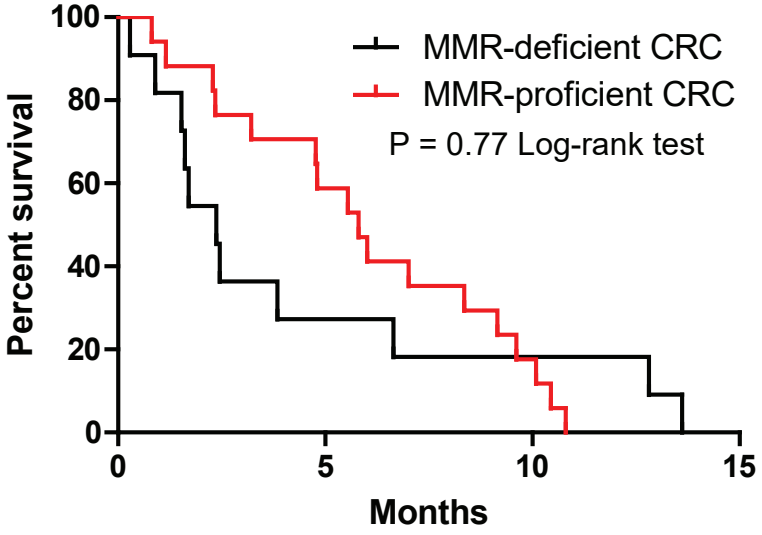


Figure S3. MMR-proficient and deficient CRCs have comparable time on treatment and duration of metastatic disease prior to study enrollment. Kaplan-Meier estimates of (A) time on therapy immediately prior to study enrollment (HR 0.81, 95% CI 0.38 to 1.752, p=.60) and (B) duration of metastatic disease prior to enrollment (HR 1.13, 95% CI 0.49 to 2.62, p=.78) on this pembrolizumab study were comparable between the MMR-deficient and proficient CRC cohorts. The short duration on prior therapy is expected in a treatment refractory CRC population.

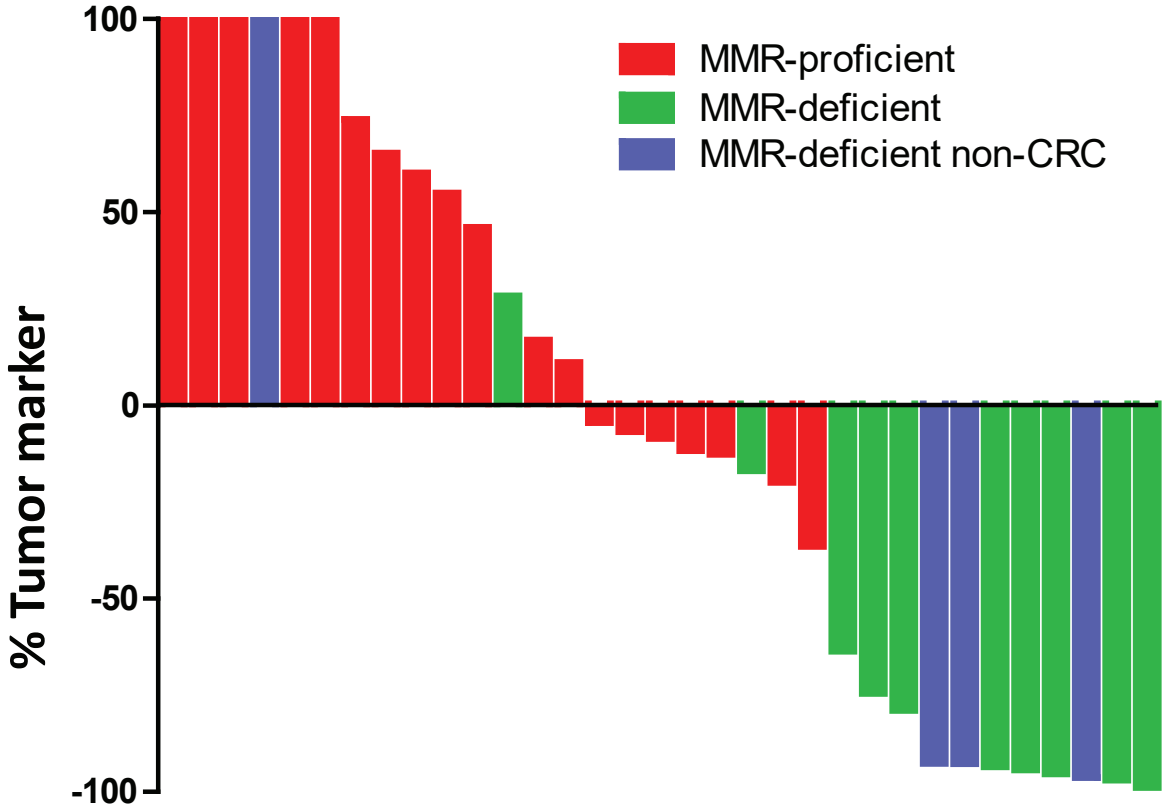


Figure S4. Waterfall plot of biochemical response. Serum protein biomarker levels were measured with each cycle and the values represent best percent change from baseline. Patients were included if baseline tumor marker values were greater than the upper limit of normal. CA-125 was used for a patient with endometrial cancer; CA19-9 was used for 1 cholangiocarcinoma and 1 ampullary cancer; and CEA was used for all other patients. Green and red represent patients with MMR-deficient and proficient CRCs, respectively. Blue represents patients with MMR-deficient cancers other than CRC.

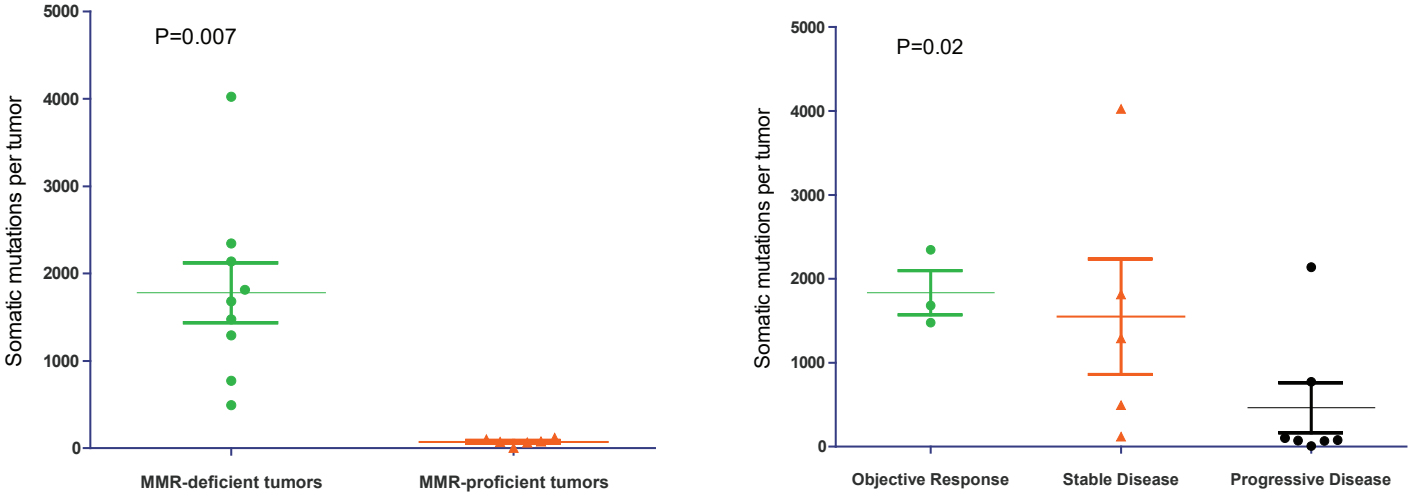


Figure S5. Somatic mutations in MMR-deficient and proficient tumors. Total somatic mutations per tumor identified by exome sequencing of tumor and matched normal DNA (left) and correlation with objective responses (right) (non-parametric Wilcoxon test , p= 0.007 and Jonckheere-Terpstra test for trend, p=0.02).

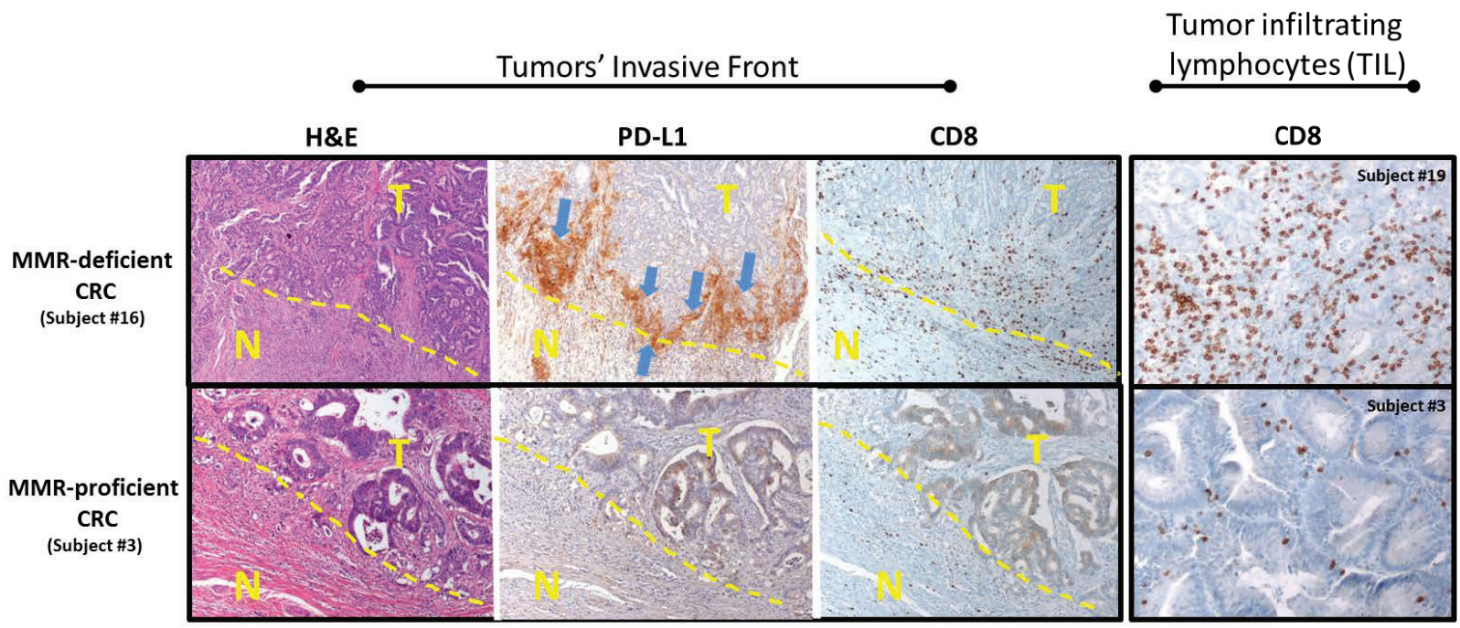


Figure S6. Immunohistochemistry of CD8 and PD-L1 Expression. The invasive front (yellow dashed line) from a MMR-deficient CRC (subject #16, top) and MMR-proficient CRC (subject #3, bottom). The yellow dashed line separates tumor (T) and normal (N) tissue. There is marked expression of PD-L1 (blue arrows) and CD8 (brown dots) in the MMR-deficient tumor (top panels) patient while there is very little expression of either marker in the MMR-proficient tumor (bottom panels). Representative images of tumor infiltrating lymphocytes (TIL) in another MMR-deficient CRC (subject #19, top) and MMR-proficient CRC (subject #3, bottom) immunolabeled with an antibody to CD8 (brown dots). Note the infiltration of CD8 cells in the MMR-deficient tumor. Invasive front original magnification 10x and TIL 20x.

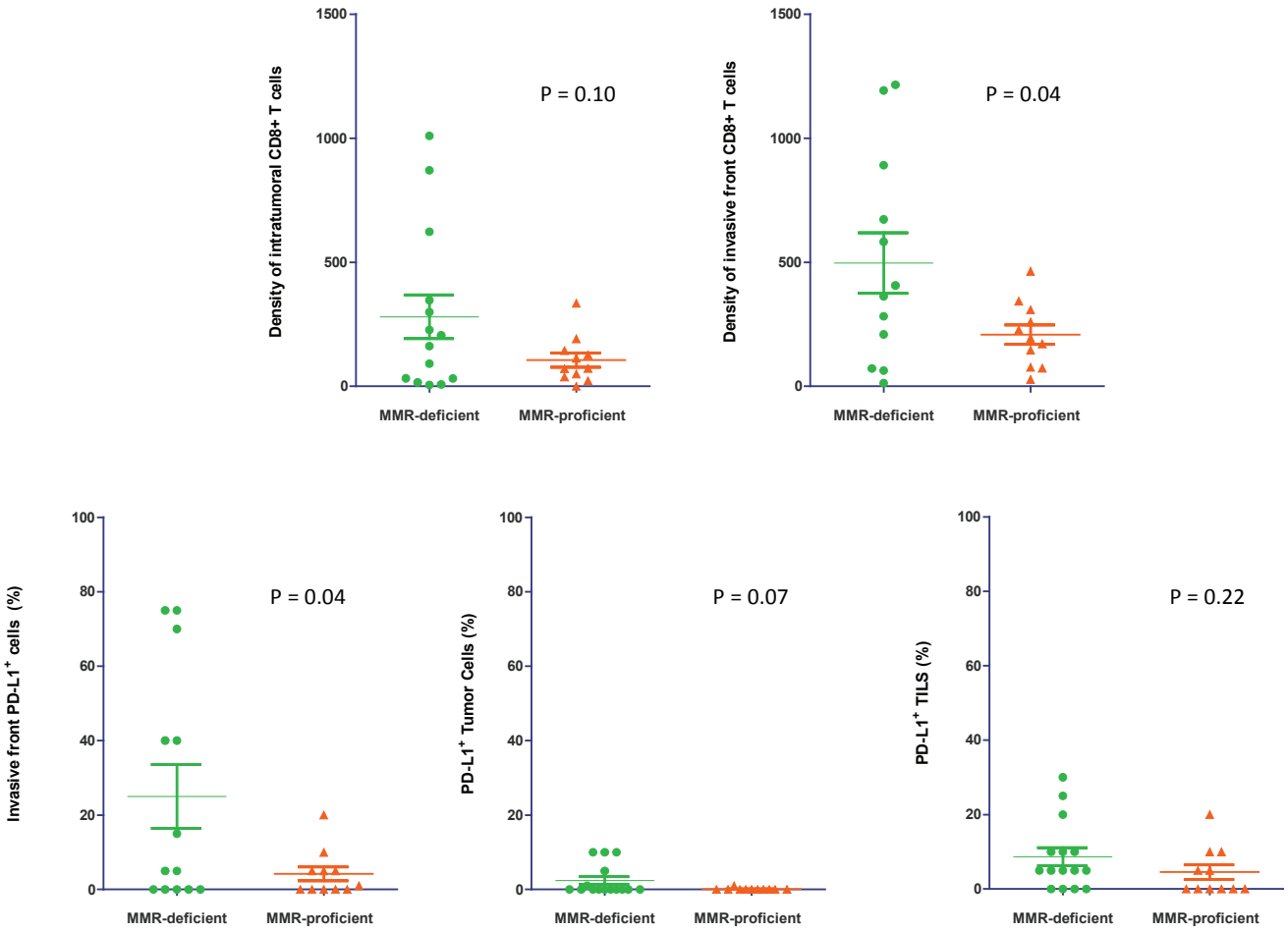
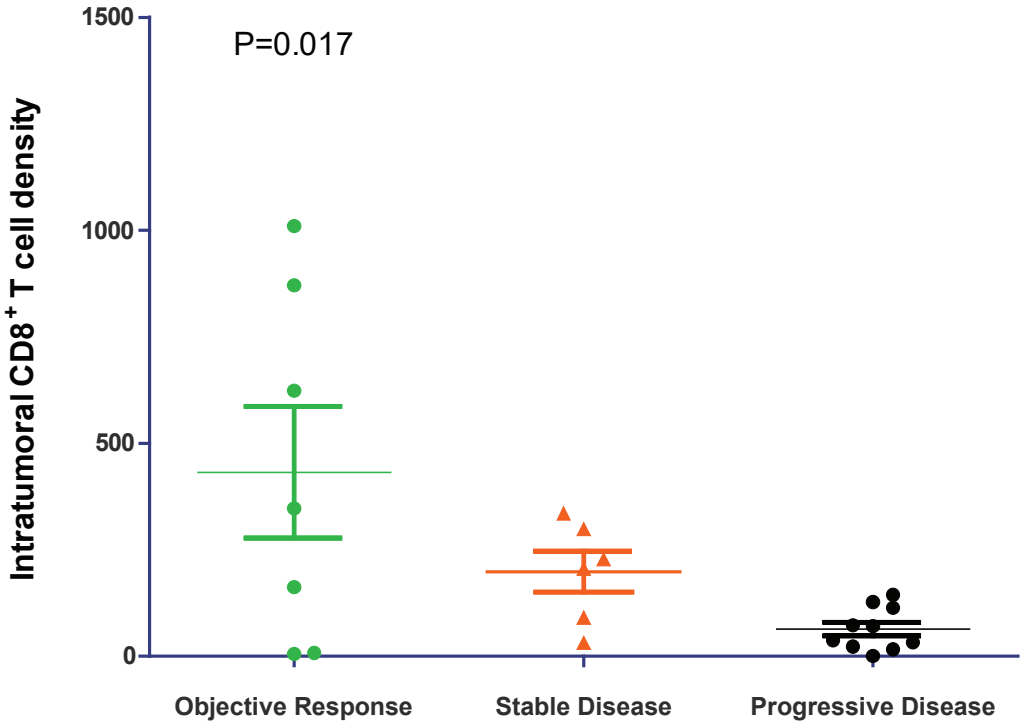


Figure S7. CD8 and PD-L1 Expression in the MMR-deficient and MMR-proficient tumor microenvironment. T cell density units are cells/mm2 of tumor. Invasive front refers to the immune cells (TILs and macrophages) at the junction of the tumor and normal tissue. P-values obtained using an unpaired t-test.



.
Figure S8. CD8 expression and clinical benefit to pembrolizumab. Correlation between the intratumoral CD8⁺ T cell density (cells/mm2) and objective response (Jonckheere-Terpstra test for trend, p=0.02).

IV. SUPPLEMENTARY TABLES

Table S1. Comparison of immune-related and RECIST response criteria
(adapted from Wolchok et al. Clin can res 2009;15:7412-20.)

	CR	PR	SD	PD
RECIST 1.1	All lesions gone	≥ 30% decrease in sum of target lesion diameters from baseline	Neither CR, PR, or PD from nadir	≥ 20% increase in sum of target lesion diameters from nadir Appearance of new lesion
	irCR	**irPR**	**irSD**	**irPD**
irRC	All lesions gone	SPD of index + any new lesions decreases ≥50% from baseline New lesions allowed	SPD of index + any new lesions neither irCR, irPR, or irPD New lesions allowed	SPD of index + any new lesions increases ≥ 25% from nadir irPD is based on SPD only

RECIST 1.1= Response Evaluation Criteria in Solid Tumors 1.1; irRC= Immune Related Response Criteria
CR= complete response; PR= paritial response; SD= stable disease; PD= progressive disease; SPD= sum of the product of the perpindicular diameters

Table S2. Immune-Related response to treatment

Type of immune-related responses-no (%)	MMR-deficient CRC n=10	MMR-proficient CRC n=18	MMR-deficient non-CRC n=7
Complete Response	0 (0)	0 (0)	1 (14)[1]
Partial Response	4 (40)	0 (0)	4 (57)[2]
Stable Disease (Week 12)	5 (50)	2 (11)	0 (0)
Progressive Disease	1 (10)	11 (61)	2 (29)
Not Evaluable[3]	0 (0)	5 (28)	0 (0)
Immune-related objective response rate (%)	40	0	71
95% CI	12-74	0-19	29-96
Immune-related disease control rate (%)[4]	90	11	71
95% CI	55-100	1-35	29-96
Immune-related PFS at 20 weeks (%)	78	11	67
95% CI	40-97	1-35	22-96

[1] Originally PR at 12 that was converted to CR at 20 weeks

[2] One PR at 12 weeks

[3] Patients were considered not evaluable if they did not undergo a 12 week scan due to clinical progression.

[4] The rate of disease control was defined as the percentage of patients who had a complete response, partial response or stable disease for 12 weeks or more.

Table S3. Summary of genomic and IHC data

See attached spreadsheet

Table S4. Correlation of total somatic mutations and mutation associated neoantigens (MANA) with clinical outcomes

Trend test to compare CR/PR vs. SD vs. PD

Variable	P value
Total Somatic Mutations	0.024
Mutation-associated neoantigens (MANAs)	0.028

Association of mutations and MANA with response rates

		95% Confidence interval		
Variable (log scale*)	Odds Ratio	Lower	Upper	P value
Total Somatic Mutations	3.312	0.500	21.924	0.214
Mutation-associated neoantigens (MANAs)	3.087	0.617	15.439	0.170

Association of mutations and MANA with PFS

		95% Confidence interval		
Variable (log scale*)	Hazard Ratio	Lower	Upper	P value
Total Somatic Mutations	0.628	0.424	0.931	0.021
Mutation-associated neoantigens (MANAs)	0.599	0.391	0.916	0.018

Association of mutations and MANA with OS

		95% Confidence interval		
Variable (log scale*)	Hazard Ratio	Lower	Upper	P value
Total Somatic Mutations	0.707	0.481	1.038	0.077
Mutation-associated neoantigens (MANAs)	0.676	0.453	1.008	0.055

Table S5. Correlation of immune markers with clinical outcome

Trend test to compare CR/PR vs. SD vs. PD

Variable	P value
PD-L1 membranous expression within tumor (≥5% vs. <5%)	0.208
PD-L1 expression by TIL and associated macrophages at invasive front (%)	0.043
PD-L1 expression by TIL and associated macrophages on TILS within tumor (%)	0.024
CD8 density (cells/mm2) within tumor	0.017
CD8 density (cells/mm2) at Invasive front	0.768
CD8 density (cells/mm2) in non-tumor tissue	0.346

* except for PD-L1 Expression on Tumor (>5% vs. <5%), which is represented as a binary variable

Association of immune markers with response rates

		95% Confidence interval		
Variable (log scale*)	Odds Ratio	Lower	Upper	P value
PD-L1 membranous expression within tumor (≥5% vs. <5%)	3.200	0.354	28.945	0.301
PD-L1 expression by TIL and associated macrophages at invasive front (%)	1.075	0.820	1.411	0.600
PD-L1 expression by TIL and associated macrophages on TILS within tumor (%)	1.139	0.862	1.506	0.361
CD8 density (cells/mm2) within tumor	1.337	0.706	2.531	0.372
CD8 density (cells/mm2) at Invasive front	0.515	0.193	1.374	0.185
CD8 density (cells/mm2) in non-tumor tissue	1.219	0.535	2.777	0.637

* except for PD-L1 Expression on Tumor (>5% vs. <5%), which is represented as a binary variable

Association of immune markers with PFS

		95% Confidence interval		
Variable (log scale*)	Hazard Ratio	Lower	Upper	P value
PD-L1 membranous expression within tumor (≥5% vs. <5%)	0.196	0.025	1.518	0.119
PD-L1 expression by TIL and associated macrophages at invasive front (%)	0.899	0.784	1.031	0.126
PD-L1 expression by TIL and associated macrophages on TILS within tumor (%)	0.888	0.765	1.031	0.118
CD8 density (cells/mm2) within tumor	0.900	0.794	1.020	0.100
CD8 density (cells/mm2) at Invasive front	1.017	0.607	1.706	0.948
CD8 density (cells/mm2) in non-tumor tissue	1.144	0.780	1.677	0.492

* except for PD-L1 Expression on Tumor (>5% vs. <5%), which is represented as a binary variable

Association of immune markers with OS

		95% Confidence interval		
Variable (log scale*)	Hazard Ratio	Lower	Upper	P value
PD-L1 membranous expression within tumor (≥5% vs. <5%)	0.513	0.062	4.232	0.535
PD-L1 expression by TIL and associated macrophages at invasive front (%)	0.978	0.816	1.173	0.813
PD-L1 expression by TIL and associated macrophages on TILS within tumor (%)	0.931	0.762	1.138	0.487
CD8 density (cells/mm2) within tumor	0.918	0.813	1.036	0.167
CD8 density (cells/mm2) at Invasive front	0.986	0.502	1.934	0.967
CD8 density (cells/mm2) in non-tumor tissue	1.018	0.635	1.631	0.941

* except for PD-L1 Expression on Tumor (>5% vs. <5%), which is represented as a binary variable

V. REFERENCES

1. Bacher JW, Flanagan LA, Smalley RL, et al. Development of a fluorescent multiplex assay for detection of MSI-High tumors. Disease markers 2004;20:237-50.
2. Murphy KM, Zhang S, Geiger T, et al. Comparison of the microsatellite instability analysis system and the Bethesda panel for the determination of microsatellite instability in colorectal cancers. The Journal of molecular diagnostics : JMD 2006;8:305-11.
3. Jones S, Anagnostou V, Lytle K, et al. Personalized genomic analyses for cancer mutation discovery and interpretation. Science translational medicine 2015;7:283ra53.
4. Sausen M, Leary RJ, Jones S, et al. Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. Nature genetics 2013;45:12-7.
5. Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. Journal of molecular biology 1970;48:443-53.
6. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsen M. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. Nucleic acids research 2008;36:W509-12.
7. Buyse M, Michiels S, Sargent DJ, Grothey A, Matheson A, de Gramont A. Integrating biomarkers in clinical trials. Expert review of molecular diagnostics 2011;11:171-82.
8. Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clinical cancer research : an official journal of the American Association for Cancer Research 2009;15:7412-20.
9. Llosa NJ, Cruise M, Tam A, et al. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. Cancer Discov 2015:43-51.
10. Taube JM, Anders RA, Young GD, et al. Colocalization of Inflammatory Response with B7-H1 Expression in Human Melanocytic Lesions Supports an Adaptive Resistance Mechanism of Immune Escape. Science Translational Medicine 2012;4:127ra37.
11. Cuka N, Hempel H, Sfanos K, De Marzo A, Cornish T. PIP: An Open Source Framework for Multithreaded Image Analysis of Whole Slide Images. LABORATORY INVESTIGATION 2014;94:398A-A.
12. Cupitt J, Martinez K. VIPS: an image processing system for large images. Electronic Imaging: Science & Technology; 1996: International Society for Optics and Photonics. p. 19-28.