Global pattern of CD8+ T cell infiltration and exhaustion in colorectal cancer predicts cancer immunotherapy response

Sun Tian (suntian.gw@gmail.com)
Carbon Logic Biotech (HK) Limited, Hongkong, China
https://orcid.org/0000-0002-0133-4454

Fulong Wang
Department of colorectal cancer, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Guangzhou 510060, China

Rongxin Zhang
Department of colorectal cancer, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Guangzhou 510060, China

Gong Chen (chengong@sysucc.org.cn)
Department of colorectal cancer, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Guangzhou 510060, China

Research Article

Keywords: biomarker, colorectal cancer, cancer immunotherapy response

DOI: https://doi.org/10.21203/rs.3.rs-89397/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

MSI/MSS status does not fully explain cancer immunotherapy response in colorectal cancer. We used gene expression data of 454 samples (MSI=131, MSI-L=23, MSS=284, Unknown=16) and developed a method TMEPRE that models colorectal cancer specific signature of CD8+ T cell infiltration and CD8+ T cell exhaustion states. TMEPRE showed predictive power in three datasets of anti-PD1 treated patients (p=0.056, 0.115, 0.003). CD8+ T cell exhaustion component of TMEPRE model correlates with anti-PD1 responding progenitor exhausted CD8+ T cells in both tumor and viral infection (p=0.048, 0.001). Global pattern of TMEPRE on 454 colorectal cancer samples indicated that 10.6% of MSS patients and 67.2% of MSI patients show biological characters that can benefit from anti-PD1 treatment. Within MSI nonresponders, approximately 50% showed no sufficient amount of tumor infiltrating CD8+ T cells and 50% showed terminal exhaustion of CD8+ T cells. These terminally exhausted CD8+ T cells coexisted with signature of myeloid-derived suppressor cells in colorectal cancer.

Introduction

Immune checkpoint inhibitors produce durable responses in some microsatellite unstable (MSI) colorectal cancer patients. But still, approximately 60% of MSI colorectal cancer patients do not respond to single immune checkpoint inhibitor treatment such as anti-PD1, and approximately 40% of MSI colorectal patients do not respond to combinations of immune checkpoint inhibitors treatment.\(^1\) The mechanism of resistance remains to be unclear. In colorectal cancer, MSI/MSS status is widely used as an indication whether a patient should receive immunotherapy. Much of the studies in colorectal cancer were therefore focused on the comparison between MSI (immune hot) tumors and MSS (immune cold) tumors, while these studies produced insights on the difference between these two colorectal cancer subtypes, they do not explain why resistance to immune checkpoint inhibitor treatment occurs within MSI colorectal tumor. In addition to MSI/MSS status, other biomarkers such as TMB, PDL1, POLE/POLD1 mutation, or MSI-like gene signature are also being used in colorectal cancer.\(^2,3\) Essentially, PDL1 provides a direct indication whether a tumor sample of a colorectal cancer patient has high CD8+ T cell infiltration, while MSI/MSS status, TMB, POLE/POLD1 mutation or MSI-like gene signature characterize the likelihood of a tumor sample generating high neoantigen level, thus provides an indirect indication whether a tumor sample of a colorectal cancer patient could potentially have high CD8+ T cell infiltration. However, it is already evident from studies of responses to anti-PD1 in lung cancer and melanoma that the number of tumor infiltrating CD8+ T cells is not the only requirement of response to anti-PD1 treatment, the characters of exhaustion state of tumor infiltrating CD8+ T cells is also required.\(^4 \sim 6\) Therefore, regardless of how technically robust biomarkers such as MSI/MSS status, TMB, PDL1, POLE/POLD1 mutation, and MSI-like gene signature can achieve, these biomarkers alone will never fully explain anti-PD1 resistance in colorectal cancer.

As melanoma is a main prototype of immune hot tumor that is used to study cancer immunotherapy, to date, most gene expression based prediction methods of cancer immunotherapy response were
developed for melanoma patients. However, these methods do not necessarily reflect tumor microenvironments where CD8+ T cells reside in colorectal cancer in which the typical immune hot population is only about 25%, there is a lack of prediction method specifically designed for colorectal cancer. The mechanism of how anti-PD1 works is still not fully understood. Studies showed anti-PD1 can directly reinvigorate exhausted CD8+ T cells and also indirectly recruiting new CD8+ T cells from peripheral sites. However, a tumor at least has two well documented primary immune escape mechanisms to become resistant to anti-PD1 treatment: lack of CD8+ T cell infiltration and CD8+ T cell dysfunction. Therefore, a colorectal tumor should meet at least two steps to become a responder to anti-PD1 treatment. Firstly, an anti-PD1 treatment responding tumors should have CD8+ T cell infiltration. Secondly, at least a subset of tumor infiltrating CD8+ T cells display properties that can respond to anti-PD1. In this report, we aim to dissect gene expression patterns of these two characters from tumor microenvironment in colorectal cancer, and develop a colorectal cancer specific method TMEPRE for the prediction of anti-PD1 response.

Results

TMEPRE model predicts anti-PD1 treatment response

The model was developed using gene expression data of colorectal cancer patients and has two components: (28 genes, Supplementary table 1) and (29 genes, Supplementary table 2).

To date, most of available gene expression datasets of anti-PD1 response were performed using melanoma as the model system. TMEPRE was validated on 3 datasets of melanoma patients who received anti-PD1 treatment. In the first dataset, the survival analysis of the TMEPRE prediction model resulted in a significant hazard ratio (n=16, pretreatment samples, GSE78220, HR=4.59, p-value=0.056, Figure 1A). In the second dataset, although the p-value of the survival analysis of the TMEPRE prediction model is large (n=21, sampling before cycle 1 day 0, GSE91061, HR=2.12, p-value=0.115, Figure 1B), the separation of survival between TMEPRE predicted responder group and TMEPRE predicted nonresponder group is still clearly observed. In the third dataset, the survival analysis of the TMEPRE prediction model resulted in a significant hazard ratio (n=21, sampling at an early treatment time point before cycle 1 day 29, GSE91061, HR=5.04, p-value=0.003, Figure 1C). Data used to validate anti-PD1 response (GSE78220 and GSE91061) are from melanoma patients and on the Illumina HiSeq platform. Data used to train the model (GSE13294, GSE26682, GSE18088, GSE39084) are from colorectal cancer patients and on the Affymetrix platform. No anti-PD1 response data or survival data was used in the training of the TMEPRE model. Despite technical noise introduced by different cancer types and different data platforms, TMEPRE model showed predictive powers for responders to anti-PD1 treatment in all three validations.

The underlying biology of TMEPRE model measures amounts of tumor infiltrating CD8+ T cells and characters of tumor infiltrating terminally exhausted CD8+ T cells
In the dataset of all 454 samples, the countings of tumor infiltrating cytotoxic lymphocytes were read out using MCP-counter and TIDE cytotoxic T lymphocytes count. The first component of TMEPRE model, score, positively correlates with counting of MCP-counter cytotoxic lymphocytes (r=0.82, r.msi=0.81) and TIDE cytotoxic T lymphocytes (r=0.68, r.msi=0.83) (Figure 2). The relative ranges coverage of scores of MSS samples is larger than the relative range coverage of MCP-counter score and TIDE score(=0.89, =0.67, =0.81). These results suggested that in MSS colorectal tumors that harbor fewer tumor infiltrating immune cells, might be a more sensitive measurement of tumor infiltrating cytotoxic lymphocytes. The reason might be that is specifically designed for the tumor microenvironment of colorectal cancer, whereas the other methods like MCP-counter and TIDE are not.

The second component of TMEPRE model, score, is designed to measure whether tumor infiltrating CD8+ T cells can respond to anti-PD1 treatment. To test whether indeed capture this character of tumor infiltrating CD8+ T cells, we read out the scores of signature in two subgroups of dysfunction CD8+ T cells isolated from tumors and chronic viral infection: terminally exhausted tumor infiltrating CD8+ T cells that can no longer respond to anti-PD-1 therapy and progenitor exhausted tumor infiltrating CD8+ T cells that can still respond to anti-PD-1 therapy (GSE122713). Because signature is derived from gene expression data of bulk tumor sample, the source of gene expressions originates from a mixture CD8+ T cells, tumor cells and other tumor infiltrating immune cell types in the tumor microenvironment, while the progenitor/terminal exhausted tumor infiltrating CD8+ T cells data are generated from isolated CD8+ T cells. Therefore, when scores were read out, only genes in that primarily originated from CD8+ T cells are used. For each of 29 genes in , median expression values of 16 purified main immune cell types were compared using BloodSpot with HemaExporer human hematopoiesis database. A gene is considered as mainly expressed by CD8+ T cells when CD8+ T cell is among the top 2 immune cell types expressing this gene. Seven genes (CCL5, CD2, CD48, CD84, FAM78A, HCST, IL21R) in passed these criteria and two genes in are inhibitor receptors on CD8+ T cells used to define early terminal exhausted CD8+ T cells (HAVCR2, PDCD1). These nine genes were used to read out scores in the isolated progenitor/terminal exhausted tumor infiltrating CD8+ T cells dataset. In both tumors and chronic viral infection, the score of are significantly higher in the subgroup of progenitor exhausted tumor infiltrating CD8+ T cells (<0.001, = 0.048, Figure 3). Therefore, the score of indeed captures the characters of progenitor exhausted tumor infiltrating CD8+ T cells that can still respond to anti-PD1.

Global pattern of TMEPRE model in MSI and MSS colorectal tumors

The TMEPRE model was read out in 454 colorectal samples (MSI=131, MSI-L=23, MSS=284, Unknown=16). A splitted heatmap was plotted. Tumors displaying a pattern of sufficient CD8+ T cell infiltration but no pattern of CD8+ T cell terminal exhaustion are considered as potential responders to anti-PD1 therapy (Figure 4).

Within 284 MSS tumor samples, 10.6% (n=30) are classified as responders and 89.4% (n=254) as nonresponders. Among MSS nonresponders, 86.6% (n=246) showed no sufficient tumor infiltrating CD8+ T cell, 2.8% (n=8) showed sufficient tumor infiltrating CD8+ T cell but those CD8+ T cells display patterns.
of terminally exhausted CD8+ T cell. As expected, the anti-PD1 resistance mechanism of the majority of MSS tumors is no sufficient amount of tumor infiltrating CD8+ T cells.

Within 131 MSI tumor samples, 67.2% (n=88) are classified as responders and 32.8% (n=43) as nonresponders. Among the MSI nonresponders, 16.0% (n=21) showed no sufficient tumor infiltrating CD8+ T cell, 16.8% (n=22) showed sufficient tumor infiltrating CD8+ T cell but those CD8+ T cells display patterns of terminally exhausted CD8+ T cells. Therefore, approximately 50% of MSI nonresponders are caused by terminal exhaustion of CD8+ T cells in the tumor microenvironment, and the rest 50% of MSI nonresponders are caused by no sufficient amount of tumor infiltrating CD8+ T cells.

TMEPRE model identified 10.6% of MSS and 67.2% of MSI colorectal cancer patients whose tumors show biological characters that can potentially benefit from anti-PD1 treatment. These predicted percentages of responders in MSS tumors and MSI tumors are consistent with the reported benefit of immune related disease control rate at 20 weeks of a cohort of colorectal cancer patients treated with pembrolizumab.\textsuperscript{16}

**Discussion**

We developed a computational method TMEPRE for colorectal cancer patients, which measures both CD8+ T cell infiltration() and whether tumor infiltrating CD8+ T cells can respond to cancer immunotherapy(). TMEPRE was developed without using any response data or survival data and was specifically designed to reflect biology of the tumor microenvironment of colorectal cancer. The method was validated in three anti-PD1 treated datasets.

Another example of a prediction method of anti-PD1 response using tumor microenvironment of CD8+ T cell exclusion and CD8+ T cell exhaustion is TIDE method.\textsuperscript{8} TIDE method also has good validation performance in anti-PD1 treated melanoma data. TIDE method was trained using survival data of melanoma and was specifically designed for five cancer types including melanoma (melanoma, neuroblastoma, triple-negative breast cancer, endometrial cancer and acute myeloid leukemia). On the contrary, TMEPRE method is specifically designed for colorectal cancer. Partially due to the methods are optimized for different cancer types, the overlap between genes used in the TMEPRE model and genes used in the TIDE model are only two(IL21R, GZMA). We expect prediction scores of TMEPRE will be a more sensitive and specific reflection of the tumor microenvironment of colorectal cancer because some characters trained in other cancer types may not be applicable in colorectal cancer. One notable example is transcriptional factor TCF7 that is an established marker in melanoma. In a dataset of 16291 immune cells from 48 tumor samples of melanoma patients treated checkpoint inhibitors, TCF7+CD8+ T cells define memory-like states and predict clinical response to checkpoint inhibitors.\textsuperscript{5} However, in colorectal cancer, the expression level of TCF7 in MSS tumors is significantly higher than MSI tumors (p-value<0.001, Figure 5). The reason might be that cell types other than CD8+ T cells express a high amount of TCF7 in the tumor microenvironment in colorectal cancer. The exact cell types are not clear, but at least, this result indicated as long as gene expression data from bulk tumor sample is used, using
expression level of TCF7 as a marker to predict whether a tumor sample will respond to checkpoint inhibitor will be technically difficult to implement in colorectal cancer.

The second component of TMEPRE, , measures whether tumor infiltrating CD8+ T cells can still respond to checkpoint inhibitors. It should be noted that because gene expression data were generated using bulk tumor samples, genes listed in are expressed on both CD8+ T cells and other immune cell types in the tumor microenvironment. Among the genes in, two genes are known inhibitor receptors on CD8+ T cells (HAVCR2, PDCD1) and seven genes (CCL5, CD2, CD48, CD84, FAM78A, HCST, IL21R) have high expression levels in purified CD8+ T cells. The expression values of these nine genes are higher in nonresponders and this correlates with the terminally exhausted type of CD8+ T cells. Among the other genes in, CIQB, CIQC, KMO, FCGR1A, FCGR1B, FCER1G show higher expression in nonresponders and these are markers of myeloid derived suppressor cells. This pattern suggests that the reason why terminally exhausted CD8+ T cell failed to respond anti-PD1 therapy might not only be the co-expression of multiple inhibitors on CD8+ T cells themselves, but also the tumor microenvironment of colorectal tumor in which those terminally exhausted CD8+ T cells located may be infiltrated with myeloid derived suppressor cells. Approximately 50% of the MSI nonresponders have high scores but have low scores, for those colorectal cancer patients, a combination of anti-PD1 with drugs targeting myeloid derived suppressor cells or a combination of drugs targeting other co-expressed inhibitors could be considered.17

By only assessing MSI/MSS status, MSS colorectal cancer patients are not recommended to be treated with anti-PD1. However, data from clinical trials showed that the disease control rate in pembrolizumab treated metastatic MSS colorectal cancers was 11%, and further, in a more recent clinical trial of neoadjuvant setting, the pathological response rate of ipilimumab + nivolumab treated early-stage MSS colorectal cancers is 27%.16,18 These results indicated responders to anti-PD1 treatment exist within the MSS colorectal cancer population. In our analysis, approximately 10.6% of MSS tumor samples showed both high scores and high scores, suggesting the biological characters of tumor microenvironments of this 10.6% of MSS patients can still potentially benefit from anti-PD1 treatment. As the number of MSS patients is much larger than MSI patients, in this dataset of 451 patients used for this study, 10.6% of MSS patient means patients should be considered for anti-PD1 treatment increase 23%. Although this prediction agrees with the reported immune related progression survival rate of MSS patients treated with pembrolizumab, the clinical utility of TMEPRE model remains to be further validated at our cancer center.

To conclude, we develop a colorectal cancer specific method TMEPRE that predicts cancer immunotherapy response. The global patterns of TMEPRE in colorectal cancer patients explained the mechanism underlying the response of anti-PD1 in MSS patients and the resistance of anti-PD1 in MSI patients. TMEPRE will aid personalized medicine options of cancer immunotherapy for colorectal cancer patients.

Methods
**Data used for the development of TMEPRE model**

Publicly available gene expression data with known MSI/MSS status of four colorectal cancer datasets (GSE13294\textsuperscript{19}, GSE26682\textsuperscript{20}, GSE18088\textsuperscript{21}, GSE39084\textsuperscript{22}) were downloaded from GEO database. All four datasets are from the same Affymetrix Human Genome U133 Plus 2.0 Array platform and normalization was performed using the frozen RMA (fRMA) method in \textit{frma} package.\textsuperscript{23} The batch effects of samples in four datasets were removed using \textit{ComBat}.\textsuperscript{24} In total, gene expression data of 454 samples were collected (MSI=131, MSI-L=23, MSS=284, Unknown=16).

**Data used for testing predictive values of the anti-PD1 response of TMEPRE model**

To test the predictive value of anti-PD1 response, three RNAseq datasets were downloaded from GEO database. The first dataset includes normalized RNAseq data and clinical data of pretreatment samples of melanoma patients who received pembrolizumab or nivolumab. Patients in this cohort who received MAPK inhibitor were removed (n=16, GSE78220).\textsuperscript{25} The second dataset includes normalized RNAseq data and clinical data of samples of melanoma patients who received nivolumab. Samples at the early treatment time point before cycle 1 day 29 and samples at the pretreatment time point before cycle 1 day 0 were analyzed separately. Patients who received priori ipilimumab treatment or with incomplete overall survival data were removed (n=21, GSE91061).\textsuperscript{26} The third dataset includes normalized RNAseq data from progenitor exhausted and terminally exhausted CD8+ T cells isolated from tumors and chronic viral infection (n=20, GSE122713).\textsuperscript{13}

**Design of TMEPRE model**

The score function of the TMEPRE model comprises of two components: TME1.TcellInfiltration and TME2.TcellResponse.

(1) TME1.TcellInfiltration scores tumor microenvironment that allows CD8+ T cell infiltration. To estimate the abundance of CD8+ T cells, we use the expression level of CD8A. The cutoff of CD8A level is defined as 40% percentile of CD8A expression level in 131 MSI tumors. MSI tumors with CD8A level higher than cutoff is defined as tumors with high CD8+ T cell infiltration (n=78); MSS tumors with CD8A level lower than cutoff is defined as tumors without high CD8+ T cell infiltration (n=211). 200 rounds of 10-fold cross-validation between these two groups were performed. In each cross-validation round, a t-test for each gene was performed and p-values of genes were ranked. CD8A gene itself was excluded from the cross-validation procedure. Genes with p-values ranked in the top 60 ranking genes in at least 80% of 200 rounds of cross-validations were selected as the signature of TME1.TcellInfiltration. The relative range coverage $\omega$ of TME1.TcellInfiltration scores of MSS samples is defined as

$$\omega = (\max \text{. score}_{mss} - \min \text{. score}_{mss}) / (\max \text{. score}_{all} - \min \text{. score}_{all})$$
(2) TME2.TcellResponse scores tumor microenvironment that tumor infiltrating CD8+ T cells do not display a terminal exhaustion pattern and can still respond to checkpoint inhibitors. To define the terminal exhaustion pattern of tumor infiltrating CD8+ T cells, we use the pattern of co-expression of multiple inhibitory receptors of PD1 and TIM3 because TIM3 is an early acquired co-expressed inhibitor receptor among all co-expressed inhibitory receptors.\(^6\) Within MSI tumors with high CD8+ T cell infiltration defined in the previous step (n=78), the median of PD1 expression level is used as the cutoff of PD1, and the median of TIM3 expression level is used as the cutoff of TIM3. MSI tumors with high CD8+ T cell infiltration and both PD1 and TIM3 higher than cutoffs are defined as a tumor microenvironment of co-expression of multiple early inhibitory receptors. CD8+ T cells resided in this type of tumor microenvironment are beginning to become terminally exhausted and resist to anti-PD1 treatment (n=21). MSI tumors with high CD8+ T cell infiltration but both PD1 and TIM3 lower than cutoffs are defined as a tumor microenvironment in which CD8+ T cells can still respond to anti-PD1 treatment (n=21). 200 rounds of 10-fold cross-validation between these two groups were performed. In each cross-validation round, a t-test for each gene was performed and p-values of genes were ranked. Genes with p-values ranked in the top 60 ranking genes in at least 80% of 200 rounds of cross-validations were selected as the signature of TME2.TcellResponse score.

A colorectal tumor with either (1) low TME1.TcellInfiltration score, (2) or high TME1.TcellInfiltration score (cutoff 75% percentile) but low TME2.TcellResponse score is considered as anti-PD1 nonresponder.

**Declarations**

**Data Availability**

The list of genes in signatures and their status of high expression level on CD8+ T cells are available in supplementary tables. Data generated for the current study are available from the corresponding author on reasonable request.

**Author contributions**

Design and concept: ST, GC. Data analysis: ST. Interpretation of results: ST, GC. Resource and clinical data interpretation: FW, RZ, GC. Write the first draft: ST. Read and review the final paper: ST, FW, RZ, GC.

**Competing interests**

Employment and stocks of Carbon Logic Biotech: ST. Inventor of a patent application: ST. All remaining authors have declared no conflicts of interest.

**Ethics approval and consent to participate**

No human or animal ethics approval was required for this study.

**Transcript Profiling data**
References

1. Oliveira, A. F., Bretes, L. & Furtado, I. Review of PD-1/PD-L1 Inhibitors in Metastatic dMMR/MSI-H Colorectal Cancer. *Front. Oncol.* **9**, (2019).

2. Havel, J. J., Chowell, D. & Chan, T. A. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nat. Rev. Cancer* **19**, 133–150 (2019).

3. Tian, S. *et al.* A robust genomic signature for the detection of colorectal cancer patients with microsatellite instability phenotype and high mutation frequency. *J Pathol* (2012).

4. Li, H. *et al.* Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. *Cell* **176**, 775-789.e18 (2019).

5. Sade-Feldman, M. *et al.* Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. *Cell* **176**, 404 (2019).

6. Thommen, D. S. *et al.* Progression of Lung Cancer Is Associated with Increased Dysfunction of T Cells Defined by Coexpression of Multiple Inhibitory Receptors. *Cancer Immunol. Res.* **3**, 1344–1355 (2015).

7. Auslander, N. *et al.* Robust prediction of response to immune checkpoint blockade therapy in metastatic melanoma. *Nat. Med.* **24**, 1545–1549 (2018).

8. Jiang, P. *et al.* Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* **24**, 1550–1558 (2018).

9. Wu, T. D. *et al.* Peripheral T cell expansion predicts tumour infiltration and clinical response. *Nature* **579**, 274–278 (2020).

10. Yost, K. E. *et al.* Clonal replacement of tumor-specific T cells following PD-1 blockade. *Nat. Med.* **25**, 1251–1259 (2019).

11. Sharma, P., Hu-Lieskovan, S., Wargo, J. A. & Ribas, A. Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell* **168**, 707–723 (2017).

12. Becht, E. *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **17**, 218 (2016).

13. Miller, B. C. *et al.* Subsets of exhausted CD8+ T cells differentially mediate tumor control and respond to checkpoint blockade. *Nat. Immunol.* **20**, 326–336 (2019).

14. Bagger, F. O., Kinalis, S. & Rapin, N. BloodSpot: a database of healthy and malignant haematopoiesis updated with purified and single cell mRNA sequencing profiles. *Nucleic Acids Res.* **47**, D881–D885 (2019).

15. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinforma. Oxf. Engl.* **32**, 2847–2849 (2016).

16. Le, D. T. *et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* **372**, 2509–2520 (2015).
17. Wesolowski, R., Markowitz, J. & Carson, W. E. Myeloid derived suppressor cells - a new therapeutic target in the treatment of cancer. *J. Immunother. Cancer* **1**, 10 (2013).

18. Chalabi, M. *et al.* Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nat. Med.* **26**, 566–576 (2020).

19. Jorissen, R. N. *et al.* DNA copy-number alterations underlie gene expression differences between microsatellite stable and unstable colorectal cancers. *ClinCancer Res* **14**, 8061–8069 (2008).

20. Vilar, E. *et al.* MRE11 deficiency increases sensitivity to poly(ADP-ribose) polymerase inhibition in microsatellite unstable colorectal cancers. *Cancer Res.* **71**, 2632–2642 (2011).

21. Gröne, J. *et al.* Molecular profiles and clinical outcome of stage UICC II colon cancer patients. *Int. J. Colorectal Dis.* **26**, 847–858 (2011).

22. Kirzin, S. *et al.* Sporadic early-onset colorectal cancer is a specific sub-type of cancer: a morphological, molecular and genetics study. *PloS One* **9**, e103159 (2014).

23. McCall, M. N., Bolstad, B. M. & Irizarry, R. A. Frozen robust multiarray analysis (fRMA). *Biostat. Oxf. Engl.* **11**, 242–253 (2010).

24. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostat. Oxf. Engl.* **8**, 118–127 (2007).

25. Hugo, W. *et al.* Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell* **165**, 35–44 (2016).

26. Riaz, N. *et al.* Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell* **171**, 934-949.e16 (2017).