



**Prediction of Response to Immune Checkpoint Blockade in
Patients with Metastatic Colorectal Cancer with
Microsatellite Instability**

Journal:	<i>Annals of Oncology</i>
Manuscript ID	ANNONC-2023-0129
Article Type:	Original Article
Date Submitted by the Author:	06-Feb-2023
Complete List of Authors:	<p>Ratovomanana, Toky; UMRS 938 - CRSA, Oncology Department Nicolle, Rémy; 2- Université Paris Cité, Centre de Recherche sur l'Inflammation, Oncology Department Cohen, Romain; Sorbonne Université, medical oncology Diehl, Aurélien; 1- Sorbonne Université, INSERM, Unité Mixte de Recherche Scientifique 938 and SIRIC CURAMUS, Oncology Department Siret, Aurélie; UMRS 938 - CRSA, Oncology Department Letourneur, Quentin; UMRS 938 - CRSA, Oncology Department Buhard, Olivier; INSERM, U938 Perrier, Alexandre; UMRS 938 - CRSA, Oncology Department Guillerm, Erell; INSERM, U938 Coulet, Florence; UMRS 938 - CRSA, Oncology Department Cervera, Pascale; UMRS 938 - CRSA, Oncology Department Benusiglio, Patrick; GH Pitié-Salpêtrière APHP, UF d'Oncogénétique; Sorbonne Université, medical oncology Labreche, Karim; UMRS 938 - CRSA, Oncology Department Colle, Raphael; Hospital Saint-Antoine, Department of Medical Oncology Collura, Ada; INSERM, U938 Despras, Emmanuelle; UMRS 938 - CRSA, Oncology Department Le Rouzic, Philippe; UMRS 938 - CRSA, Oncology Department Renaud, Florence; UMRS 938 - CRSA, Oncology Department Cros, Jerome; Hopital Bichat - Claude-Bernard, Pathology Alentorn, Agusti; Assistance Publique - Hopitaux de Paris, Service Neurologie -2 Touat, Mehdi; UMRS 938 - CRSA, Oncology Department Ayadi, Mira; Ligue Nationale Contre Le Cancer, CIT Bourgoin, Pierre; Sorbonne Université, medical oncology Prunier, Céline; UMRS 938 - CRSA, Oncology Department TOURNIGAND, Christophe; AP-HP, Henri-Mondor University Hospital,</p>

	<p>Department of Medical Oncology de la Fouchardière, Christelle; Hospices Civils de Lyon-Centre Anti-Cancéreux Léon-Bérard, Consortium Cancer Thyroïdien; Centre Léon Bérard, Medical Oncology</p> <p>tougeron, david; Poitiers university hospital, gastroenterology Jonchere, Vincent; UMRS 938 - CRSA, Oncology Department</p> <p>Bennouna, Jaafar; Institut de Cancerologie de l'Ouest, Oncology de Reyniès, Aurélien; Cordeliers Research Centre, Cancer, Immune control, and Escape</p> <p>Flejou, Jean-François; UMRS 938 - CRSA, Oncology Department SVRCEK, Magali; INSERM U938, APHP Hôpital Saint Antoine; AP-HP, Hôpital Saint-Antoine, Pathology</p> <p>André, Thierry; Sorbonne Universités, Hôpital Saint Antoine, Oncologie Médicale; Sorbonne Université</p> <p>Duval, Alex; UMRS 938 - CRSA, Oncology Department</p>
Keywords:	<p>Metastatic Colorectal Cancer, Microsatellite Instability, Immune Checkpoint Inhibitors (ICI), DNA and RNA sequencing, prediction of MSI mCRC patient response to ICI, progression-free survival by iRECIST</p>
Abstract:	<p>Background: Mismatch repair deficient (dMMR) tumors displaying microsatellite instability (MSI) represent a paradigm for the success of immune checkpoint inhibitor (ICI)-based immunotherapy, particularly in patients with metastatic colorectal cancer (mCRC). However, a proportion of patients with dMMR/MSI mCRC exhibit resistance to ICI. Identification of tools predicting MSI mCRC patient response to ICI are required for the design of future strategies further improving this therapy.</p> <p>Patients and Methods: Here we combined high-throughput DNA and RNA sequencing of tumors from 116 patients with MSI mCRC treated with anti-PD-1 +/- anti-CTLA-4 of the NIPICOL phase II trial (C1, NCT03350126, discovery set) and the IMMUNOMSI prospective cohort (C2, validation set). Were clinically relevant the DNA/RNA predictors whose status was significantly associated with ICI response in both C1 and C2. Primary endpoint was iPFS (progression-free survival by iRECIST).</p> <p>Results: Analyses failed to validate previously suggested DNA/RNA indicators of resistance to ICI, e.g. MSISensor score, tumor mutational burden, or specific cellular and molecular tumoral contingents. By contrast, iPFS under ICI was shown in C1 and C2 to depend both on a multiplex MSI signature involving the mutations of 19 microsatellites (HRC2 = 13.1; 95% CI (2.71-63.8) ; p = 1.4x10⁻³) and the expression of a set of 182 RNA markers with a non-epithelial TGFB-related desmoplastic orientation (HRC2 = 4.12 ; 95% CI (1.44-11.8) ; p = 4.19x10⁻³). Both DNA (HRC2 = 6.05; 95% CI (1.68-21.8) ; p = 5.92x10⁻³) and RNA (HRC2 = 8.38 ; 95% CI (2.44-28.7) ; p = 7.22x10⁻⁴) signatures were independently predictive of iPFS.</p> <p>Conclusions: iPFS in patients with MSI mCRC can be predicted by simply analyzing the mutational status of DNA microsatellite-containing genes in epithelial tumor cells together with nonepithelial TGFB-related desmoplastic RNA markers. These DNA/RNA signatures are readily available to be used in clinics.</p>

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Prediction of Response to Immune Checkpoint Blockade in Patients with Metastatic Colorectal Cancer with Microsatellite Instability

Toky Ratovomanana ^{1€}, Rémy Nicolle ^{2,3€}, Romain Cohen ^{1,3,4}, Aurélien Diehl ¹, Aurélie Siret ¹, Quentin Letourneur ¹, Olivier Buhard ¹, Alexandre Perrier ^{1,5}, Erell Guillerm ^{1,5}, Florence Coulet ^{1,5}, Pascale Cervera ¹, Patrick Benusiglio ^{1,5}, Karim Labrèche ⁶, Raphaël Colle ^{1,3,4}, Ada Collura ¹, Emmanuelle Despras ¹, Philippe Le Rouzic ¹, Florence Renaud ¹, Jérôme Cros ⁷, Agusti Alentorn ⁸, Mehdi Touat ⁸, Mira Ayadi ⁹, Pierre Bourgoin ^{1,10}, Céline Prunier ¹¹, Christophe Tournigand ¹², Christelle de la Fouchardière ¹³, David Tougeron ¹⁴, Vincent Jonchère ¹, Jaafar Bennouna ¹⁵, Aurélien de Reynies ¹⁶, Jean-François Fléjou ^{1,10}, Magali Svrcek ^{1,10}, Thierry André ^{1,3,4}, **Alex Duval** ^{1,5\$}

- 1- Sorbonne Université, INSERM, Unité Mixte de Recherche Scientifique 938 and SIRIC CURAMUS, Centre de Recherche Saint-Antoine, Equipe Instabilité des Microsatellites et Cancer, Equipe labellisée par la Ligue Nationale contre le Cancer, F-75012 Paris, France
- 2- Université Paris Cité, Centre de Recherche sur l'Inflammation (CRI), INSERM, U1149, CNRS, ERL 8252, F-75018 Paris, France.
- 3- GERCOR, Groupe Coopérateur Multidisciplinaire en Oncologie, F-75011 Paris, France.
- 4- Sorbonne Université, Department of Medical Oncology, AP-HP, hôpital Saint-Antoine, F-75012 Paris, France.
- 5- Sorbonne Université, Department of Molecular Biology and Medical Genetics, AP-HP, Hospital Pitié-Salpêtrière, F-75012 Paris, France.
- 6- CinBioS, MS 37 PASS Production de données en Sciences de la vie et de la Santé, INSERM, Sorbonne Université et SIRIC CURAMUS, 75013 Paris.
- 7- Department of Pathology, Beaujon Hospital, AP-HP, Clichy, France.
- 8- Service de Neurologie 2-Mazarin, Sorbonne Université, Inserm, CNRS, UMR S 1127, Institut du Cerveau, ICM, AP-HP, Hôpitaux Universitaires La Pitié Salpêtrière - Charles Foix, 47-83 boulevard de l'Hôpital, 75013, Paris, France.
- 9- Programme "Cartes d'Identité des Tumeurs", Ligue Nationale Contre le Cancer, Paris, France.
- 10- Sorbonne Université, Department of Pathology, AP-HP, hôpital Saint-Antoine, F-75012 Paris, France.

- 1
2
3 11- Sorbonne Université, INSERM, Unité Mixte de Recherche Scientifique 938 and SIRIC
4 CURAMUS, Centre de Recherche Saint-Antoine, Equipe Signalisation TGFB, plasticité
5 cellulaire et Cancer, F-75012 Paris, France.
6
7 12- Department of medical Oncology, Hôpital henri Mondor, APHP, Université Paris Est Creteil,
8 INSERM U955.
9
10 13- Department of Medical Oncology, Centre Léon Bérard, Lyon, France.
11
12 14- ProDicET, UR 24144, University of Poitiers and Hepato-Gastroenterology Department, Poitiers
13 University Hospital, 86000 Poitiers, France.
14
15 15- Centre De Recherche En Cancérologie Et Immunologie Nantes-Angers (CRCINA), INSERM,
16 Université d'Angers, Université De Nantes, Nantes, France.
17
18 16- Cartes d'Identité des Tumeurs Program, Ligue Nationale Contre Cancer, Paris, France.
19

20
21 € Co-first Author

22
23 \$ Leadership, Corresponding author
24
25
26
27

28 **RUNNING HEAD:** Prediction of Response to Immune Checkpoint Blockade in
29 dMMR/MSI mCRC
30

31 **CORRESPONDENCE:** Address correspondence to: Alex Duval, Team “Microsatellite
32 Instability and Cancer”, Inserm UMRS 938, CRSA, Sorbonne University, Hôpital Saint-
33 Antoine, Paris 75010, France. e-mail: alex.duval@inserm.fr
34
35
36
37
38
39

40 **ACKNOWLEDGMENTS:** The authors thank the patients and their families for making
41 the study possible. The authors acknowledge the GERCOR clinical study teams, the
42 IDEATION study group and investigators and study teams in all centers.
43
44
45
46
47
48

49 **FUNDING:** This work was supported by grants from Site de Recherche Intégré sur le
50 Cancer (SIRIC) Cancer United Research Associating Medicine, University & Society
51 (CURAMUS), and the Ligue Nationale Contre le Cancer. T.R was a recipient of a grant
52 from the Region Ile-de-France (PhD Student bursary). Prof. Duval’s team has the label
53 de La Ligue contre le Cancer.
54
55
56
57
58
59
60

1
2
3 **AUTHOR CONTRIBUTIONS:** study concept and design: TR, RN, Adi, ADu;
4
5 acquisition of data: TR, RN, RCoh, Adi, AS, QL, OB, AP, EG, FC, PC, RCol, AC, ED,
6
7 PLR, FR, AA, MT, FB, MA, PB, CP, VJ, JB, AdR, JFF, MS, CT, CdIF, DT, TA, ADu;
8
9 analysis and interpretation of data: TR, RN, RCoh, Adi, KL, JC, ADu; drafting of the
10
11 manuscript: TR, RN, RCoh, Adi, ADu; obtained funding; ADu; technical support: TR,
12
13 RN, Adi, AS, QL, OB, FC, AC, FR, PB, VJ ; study supervision: ADu.
14

15 **DISCLOSURE:** The authors declare no conflict of interest
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

ABSTRACT

Background: Mismatch repair deficient (dMMR) tumors displaying microsatellite instability (MSI) represent a paradigm for the success of immune checkpoint inhibitor (ICI)-based immunotherapy, particularly in patients with metastatic colorectal cancer (mCRC). However, a proportion of patients with dMMR/MSI mCRC exhibit resistance to ICI. Identification of tools predicting MSI mCRC patient response to ICI are required for the design of future strategies further improving this therapy.

Patients and Methods: Here we combined high-throughput DNA and RNA sequencing of tumors from 116 patients with MSI mCRC treated with anti-PD-1 +/- anti-CTLA-4 of the NIPICOL phase II trial (C1, NCT03350126, discovery set) and the IMMUNOMSI prospective cohort (C2, validation set). We clinically relevant the DNA/RNA predictors whose status was significantly associated with ICI response in both C1 and C2. Primary endpoint was iPFS (progression-free survival by iRECIST).

Results: Analyses failed to validate previously suggested DNA/RNA indicators of resistance to ICI, e.g. MSISensor score, tumor mutational burden, or specific cellular and molecular tumoral contingents. By contrast, iPFS under ICI was shown in C1 and C2 to depend both on a multiplex MSI signature involving the mutations of 19 microsatellites ($HR_{C2} = 13.1$; 95% CI [2.71-63.8] ; $p = 1.4 \times 10^{-3}$) and the expression of a set of 182 RNA markers with a non-epithelial TGFB-related desmoplastic orientation ($HR_{C2} = 4.12$; 95% CI [1.44-11.8] ; $p = 4.19 \times 10^{-3}$). Both DNA ($HR_{C2} = 6.05$; 95% CI [1.68-21.8] ; $p = 5.92 \times 10^{-3}$) and RNA ($HR_{C2} = 8.38$; 95% CI [2.44-28.7] ; $p = 7.22 \times 10^{-4}$) signatures were independently predictive of iPFS.

Conclusions: iPFS in patients with MSI mCRC can be predicted by simply analyzing the mutational status of DNA microsatellite-containing genes in epithelial tumor cells together with nonepithelial TGFB-related desmoplastic RNA markers. These DNA/RNA signatures are readily available to be used in clinics.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Keywords: Metastatic Colorectal Cancer, Microsatellite Instability, Immune Checkpoint Inhibitors (ICI), DNA and RNA sequencing, prediction of MSI mCRC patient response to ICI, progression-free survival by iRECIST.

For Peer Review

INTRODUCTION

Mismatch repair deficient (dMMR) tumors display a molecular phenotype characterized by the genetic instability of numerous microsatellite repeated sequences throughout the genome (Microsatellite Instability, MSI) ¹⁻³. MSI was first observed in inherited tumors associated with Lynch syndrome and later in a large spectrum of primary tumors, in particular sporadic colorectal cancers (CRC) (for review see ⁴⁻⁶). Being highly genetically unstable, MSI cancers are highly immunogenic and generally show a strong infiltration with cytotoxic T-cell lymphocytes ⁴. Recently, it was reported that MSI tumors resist this hostile immune microenvironment by overexpressing immune checkpoint (ICK)-related proteins to allow immune-escape ⁷. Consistently, MSI status was shown to predict clinical benefit from ICK inhibitors (ICI) in patients with metastatic cancer including CRC (mCRC) ^{8, 9}. In patients pretreated with fluoropyrimidines, oxaliplatin and irinotecan prior to ICI for dMMR/MSI mCRC, the objective response rate ranges from 33% to 65% and the 1-year overall survival rate ranges from 34% to 71%. First-line pembrolizumab (anti-PD-1) has been associated with significant improvement of progression-free survival compared with standard of care chemotherapy ± cetuximab or bevacizumab ¹⁰. However, up to 15-46% of patients with dMMR/MSI mCRC exhibit primary resistance to ICI while 5%-25% of responders develop acquired resistance to these treatments, although this estimation might increase with longer follow-up ¹¹⁻¹⁴.

Several DNA and RNA-based markers predicting the efficacy of ICIs have been previously proposed in metastatic dMMR/MSI cancer settings ^{9, 15-21}. However, it is fair to say that these results lack independent validation, being based on the analysis of only limited series of dMMR/MSI tumor samples with heterogeneous tumor origins. At the DNA level, markers include quantitative genomic indexes measuring the level of MSI within the tumor bulk such as the MSIsensor score or the tumor mutation burden (TMB). Previous studies have considered the selection of specific somatic variants occurring in dMMR cancers due to MSI as specific predictive markers such as *BRAF*^{V600E} mutational status, *KRAS/NRAS* mutations, or the *B2M* truncating mutation resulting in loss of function of the resultant protein associated with the MHC class I complex ^{17, 20}. However, it was recently reported that B2M inactivation was unlikely to

1
2
3 blunt the efficacy of ICI in dMMR/MSI tumors in human ²² and using multiple murine
4 dMMR B2m null cancer models ²³, raising the question of the real impact of these
5 somatic mutations in the MSI cancer setting. At the RNA level, it was hypothesized
6 that the estimated abundance of specific cell populations in the tumor
7 microenvironment could be of clinical relevance, e.g., immune cell populations such as
8 antigen-presenting macrophages interacting with T-cells ²⁰. Logically, deregulated
9 activity of some cancer-related pathways associated with antitumor immunity were also
10 proposed, e.g., the reduced activity of *Wnt/Wingless* signaling, deregulation of the
11 interferon gamma pathway and/or of several immune escape processes ²⁰. Finally, a
12 study published by our team showed that the leading cause for association of primary
13 resistance to ICI in mCRC was the misdiagnosis of their MSI or dMMR status ²⁴,
14 emphasizing that the first and foremost criteria to be validated before the prescription
15 of ICI in metastatic dMMR/MSI cancer patients, and particularly in mCRC, is the
16 guarantee of a quality diagnosis with appropriate methods to identify genuine
17 dMMR/MSI.
18
19

20
21 In this study, we hypothesized that the response to ICI could be predicted by both
22 leveraging on the MSI-specific mutational patterns in epithelial cells (by DNA sequencing)
23 and by considering the modifications of the cancer stroma (by RNA sequencing) on
24 the other hand. We addressed the issue of response to ICI based on the investigation
25 of two independent prospective cohorts of 44 and 72 patients with dMMR/MSI mCRC,
26 respectively the multicentric NIPICOL clinical trial (NCT03350126) ¹⁴ and the
27 prospective ImmunoMSI cohort ²⁵.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

46 MATERIALS AND METHODS

47 Patients

48
49 The Nipicol clinical trial (C1) was conducted in accordance with the tenets of the
50 Declaration 170 of Helsinki and Good Clinical Practice Guidelines, after approval by
51 the ethics board, and was registered at ClinicalTrials.gov (Clinical trial number:
52 NCT03350126). Written informed consent was obtained from all patients. All patients
53 included in C1 received anti-PD1 (nivolumab) + anti-CTL4 (ipilimumab). Concerning
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ImmunoMSI prospective cohort (C2) ²⁵, all consecutive MSI/dMMR mCRC patients treated with ICI (anti-PD-1 monotherapy or anti-PD-1 plus anti-CTLA-4 combination) at Saint-Antoine Hospital, Medical Oncology department (Prof. Thierry André) from February 2015 to December 2019, were included. This research was approved by the ethics committee (N°2020 – CER 2020-6). The identification numbers of the ethic committees are the following, i.e., CPP N°2017/45 (ANSM reference: 170508A-12; EudraCT N°2017-002442-72) for C1 NIPICOL clinical trial and CPP N°2020 CER 2020-6C2 for ImmunoMSI. The Nipicol trial NCT03350126 has been completed and its primary endpoint published ¹⁴.

Assessment of microsatellite status

All CRC samples from C1, C2 and other cohorts used in this post-hoc study were centrally reassessed for MSI and dMMR status using immunohistochemistry (IHC) and for MSI using polymerase chain reaction (PCR) as described ^{26, 27}. Next, MSIsensor (version 0.6) and MSICare were used by default on paired normal-tumor exome data, to evaluate the mutation status of microsatellites using considering the whole exome data. MSIsensor score threshold of 10% or more was used to classify the MSI-H tumor (MSI-High) and a MSICare threshold of 20% was used to define MSI status as previously described ²⁸.

Radiological analyses

Tumors were assessed ≤ 28 days before the first dose (baseline) and every 6 to 10 weeks, thereafter, according to different protocols. The decision to pursue treatment beyond iuPD was at the discretion of the treating physician. Treatment beyond iuPD was conditional to a confirmatory imaging 4 to 8 weeks after the first evidence of progression. Imaging was retrospectively and centrally reviewed by an experienced radiologist according to RECIST1.1 and iRECIST. ²⁹. Radiological progression was defined as confirmed PD (cPD) according to iRECIST 1.1. All cases of PD and pseudo progression were reviewed by a second experienced radiologist unaware of the target and non-target lesions chosen by the first radiologist. In case of discrepancy, a final decision was reached by consensus.

Survival analysis of ICI treated patients

Survival analysis on ICI patients was performed from the date of first infusion of any ICI. iPFS by iRECIST²⁹ was calculated from the first dose to the first documented cPD, or death resulting from any cause, whichever occurred first. Kaplan Meier curves were used to visualize difference in progression free survival (iPFS) between patients' groups diverging on genomic instability (e.g., TMB-High or MSISensor-High or MSICare-High, >20th percentile; TMB-Low or MSISensor-Low or MSICare-low ≤20th percentile) or to show at the selected cutoff values a clinical effect of our DNA or RNA signatures. The cutoff values were selected to optimize the clinical effect of our DNA or RNA signatures on ICI response in patients. Nevertheless, the validity of same signatures was also tested without using cutoff values with Cox models, similarly showing a clinical relevance, as indicated in the corresponding figures. The use of the cutoffs has been envisioned more to illustrate the findings but there is no dependence on these cutoffs for the demonstration of the clinical impact of both the DNA and RNA signatures. Two-sided Log-rank test was performed using the R package Survival (version 3.2.3).

P-value levels and Correction of *p*-values

In all statistical analyses, *p*-value ≤ 0.05 (risk alpha) was used as a threshold of significance. As an exception, for the selection of DNA variants following WES in C1, *p*-value ≤ 0.1 has been applied in order to provide pre-selection of a larger number of DNA variants prior to validation of their putative clinical significance in C2 by targeted sequencing (reduction of the risk of false negatives).

We systematically applied multi-test correction using Benjamini-Hochberg's method to control the false discovery rate when multiple tests of the same hypotheses were performed. However, we chose to highlight either one or the other either for clarity or fairness of comparison. We systematically presented uncorrected and corrected *p*-values when relevant in each of the corresponding figures and supplementary figures.

Data availability

The processed datasets used in the current study are now available in the github repository, https://github.com/CRSA-MSI/R-ICI_MNB_Survival. Reasonable request for sharing biologic materials or raw data files will be reviewed by the corresponding author (ADu). Patient- related data not included in the paper were generated as part of clinical trials and may be subject to patient confidentiality. Any data and materials that can be shared will be released via a material transfer agreement.

Code availability

The scripts and the Naive Bayes model can be accessed at GitHub https://github.com/CRSA-MSI/R-ICI_MNB_Survival. Scripts for ICA of RNA can be assessed at GitHub <https://github.com/GeNeHetX/NipicolICA>.

Concerning Sample preparation and sequencing, Exome Analysis and Mutational load, MHC class-I neoepitopes prediction, Feature selection and data pre-processing, DNA signatures (mutation count and Multinomial Naives Bayes), and Transcriptome profile generation and analysis, see **SUPPLEMENTARY MATERIALS AND METHODS**.

RESULTS

Patient population and study design

A total of 129 prospectively collected mCRC patients were assessed for eligibility, including 57 patients from the NIPICOL clinical trial (C1, NCT03350126) and 72 patients from the prospective ImmunoMSI cohort (C2)²⁵. Clinical and disease characteristics of patients from C1 and C2 are summarized in **TABLE 1**, and further detailed in **SUPPLEMENTARY TABLE S1** and **SUPPLEMENTARY TABLE S2**. In C1, the analysis was restricted to 54 patients after removing 3 mCRC samples with misdiagnosed dMMR/MSI status (false positive cases that were in fact MSS). Whole Exome Sequencing (WES) and RNA Sequencing (RNASeq) were performed on 23 and 44 collected mCRC +/- matched normal colonic mucosa paraffin-embedded

1
2
3
4 samples, respectively, after removing 29 samples with lack of materials and/or
5 insufficient quality (FIG.1A). In C2, targeted next-generation-sequencing (NGS) and
6 RNASeq were performed on 66 and 72 mCRC +/- matched normal colonic mucosa
7 paraffin-embedded samples, respectively, after removing unqualified samples for
8 similar reasons, i.e., insufficient quantity and/or low-quality level (FIG.1A). The
9 CONSORT-like clinical and molecular diagram in FIGURE 1 outlines the methodology
10 workflow of the study.
11
12
13
14
15

16 FIGURE 1 also summarizes the flow chart (FIG. 1A) and the current design of the
17 study (FIG. 1B). In brief, NIPICOL (C1) was used as a training and ImmunoMSI (C2)
18 as a validation cohort. Progression-free survival (PFS) by iRECIST (RECIST for
19 Response Evaluation Criteria in Solid Tumor) (iPFS)²⁹ was used as primary endpoint.
20 In an effort to optimize the available clinical data for the purpose of identifying available
21 RNA/DNA predictive markers of resistance to ICI in this prospective series of ICI-
22 treated patients with MSI mCRC, their putative clinical relevance was examined in C1
23 (training set) and yet in C2 (validation set) independently. C1 was a cohort from a
24 clinical trial including dMMR/MSI patients all treated with the same treatment, i.e., a
25 combination of nivolumab (anti-PD-1) and ipilimumab (anti-CTLA-4) after failure of
26 standards of care. C2 was a larger and more heterogeneous cohort concerning
27 regimen of ICI received (anti-PD1 alone or associated with ipilimumab) and inclusion
28 criteria were less restrictive compared to a clinical trial.
29
30
31
32
33
34
35
36
37
38
39
40
41

42 **The level of MSI and TMB in tumor DNA does not predict response to ICI in patients** 43 **following exclusion of ICI-treated mCRC with misdiagnosed dMMR/MSI status** 44

45 We first focused DNA analyses on the genomic quantitative indexes that are related
46 to MSI in tumor DNA, i.e., the TMB and MSISensor score, whose level was previously
47 reported to predict response to ICI in patients with metastatic dMMR tumor^{15, 16}
48 (SUPPLEMENTARY FIGURE S1). Our data establish in this study that there is no
49 influence of MSISensor score nor TMB on the iPFS of patients from C1 after exclusion
50 of the cases with false positive MSI status (that were MSS, n = 3 and may lead to
51 wrongly give a prognostic value to TMB and/or MSISensor regarding iPFS, see
52 SUPPLEMENTARY FIGURE S1) (FIG. 2C). As shown also in SUPPLEMENTARY
53 FIGURE S1, the MSICare score is no more clinically relevant for predicting iPFS to ICI
54
55
56
57
58
59
60

1
2
3
4 in patients with confirmed MSI mCRC, nor is the number of neoepitopes presented by
5 autologous class I HLA alleles when calculated using a dedicated pipeline which
6 combines several software³⁰⁻³³. In view of these results, we decided not to investigate
7 further the clinical relevance of MSIsensor score, TMB and number of neoepitopes in
8 C2 since none of them were associated with patients' prognosis in a true MSI mCRC
9 population.
10
11
12
13
14
15
16

17 **Identification of a 19-plex MSI signature predictive of response to ICI in mCRC patients**

18
19 We first examined by WES the impact of somatic variants occurring in true positive
20 MSI mCRC patients from C1 regarding iPFS upon ICI treatment. Expectedly, a great
21 number of variants occurred at both nonrepetitive (NR, in $n = 3,886$ genes, only coding
22 events) or repetitive (R, in $n = 20,472$ microsatellite-containing genes, both coding or
23 noncoding events) sequences in genes having or not an expected role in the MSI-
24 driven tumorigenic process. For survival analyses, we considered only somatic
25 microsatellite mutations in tumor DNAs, because frequent hotspot NR mutations are
26 rare in these tumors (**Fig. 2A** and **2B**). However, note that canonic NR mutations
27 recurrently associated with colon tumor among which some were previously proposed
28 to affect response to ICI, e.g., in *KRAS*, *BRAF*, *B2M* and other cancer-related genes,
29 were not associated to iPFS in C1 (**SUPPLEMENTARY FIGURE S2**). In **FIGURE 2B**
30 are shown the results of univariate cox analyses we performed to identify amongst the
31 microsatellite variants the 167 alterations with impact on the iPFS of MSI mCRC
32 patients in C1 ($p \leq 0.1$).
33
34
35
36
37
38
39
40
41
42
43

44 Like the NGS-based MSICare²⁸ or the PCR Pentaplex²⁶ genomic tools we
45 previously designed, which use multiple microsatellite analysis to predict the MSI or
46 MSS status of CRC, we hypothesized that response to ICI could be also predicted by
47 investigating a set of selected microsatellite markers mutated in these tumors. From
48 C1 (Training cohort), 19 candidate microsatellite markers (i.e., in *SUCO*, *LYST*,
49 *MACO1*, *PPRC1*, *COMMD3*, *OCA2*, *MTMR10*, *IGDCC4*, *NLK*, *RNF43*, *TSEN54*,
50 *ADGRE1*, *EDAR*, *TTN-AS1*, *IQCA1*, *RWDD4*, *CCDC158*, *CANX*, *MLIP*) were selected
51 since: (i) their somatic mutations were each associated with iPFS in C1 ($\alpha = 5\%$)
52 (**Fig 2A**) and (ii) their sequencing in C1 but also in additional public (TCGA COAD) and
53 private cohorts (personal data, not shown) using NGS was successful in at least 95%
54
55
56
57
58
59
60

1
2
3
4 percent of tumors (**Fig 2A**) or in other words that these loci were particularly easy to
5 analyze. A predictive nature of this 19-plex identified DNA signature based on the
6 concomitant analysis of their microsatellite somatic variations in tumor DNA as
7 compared to matched normal DNA was then evaluated by simply calculating for each
8 patient the mean count of all the mutations in those microsatellites
9 (**SUPPLEMENTARY FIGURE S3**). We also used Multinomial Naive Bayes (MNB) as
10 a simple Machine Learning method to predict the risk of progression with the 19-plex
11 signature ³⁴. We trained the MNB method on Nipicol (C1) and validated it on
12 ImmunoMSI (C2). Applying the trained model in C2 (Validation cohort), we were able
13 to show that the same 19-plex MSI signature was significantly associated to iPFS
14 ($HR_{C2} = 13.1$; 95% CI [2.71-63.8] ; $p = 1.4 \times 10^{-3}$) (**Fig 2B**). **FIGURE 2** further illustrates
15 how the 19 microsatellite mutations contribute to predicting resistance to ICI in MSI
16 mCRC patients (see the heatmap on **FIG. 2E**). In the multivariate analysis we
17 conducted in C2 including molecular and clinical data (e.g., *KRAS* and *BRAF* mutation
18 status) as well as the type of immunotherapy, the independent prognostic value of this
19 19-plex MSI signature was confirmed ($HR_{C2} = 13.5$; 95% CI [2.38-76.6] ; $p = 3.3 \times 10^{-3}$)
20 (**Fig. 2F**).

Transcriptomic signatures and response to ICI in MSI mCRC patients

21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Previous studies have provided evidence of the predictive value of RNA signature
for immunotherapies. Three types of established signatures of response to ICI were
systematically assessed in both discovery (C1) and validation (C2) cohorts: signatures
quantifying cellular components of the tumor microenvironment (n=99), single gene
expression levels (n= 19,116), and signaling pathway activity estimation from gene
expression levels (n= 3,365). None of these established transcriptomic-based
signature could be reproducibly associated to iPFS in the two ICI-treated MSI mCRC
cohorts, including those specifically associated to ICI response in previous studies in
MSS tumors (e.g., PD-1, PD-L1 +/- CTLA-4, T or B cells) (**SUPPLEMENTARY FIGURE
S4**). More specifically, gene sets involved in angiogenesis, epithelial-to-mesenchymal
transition, canonical TGF-beta and Wnt/Wingless signaling pathways, as well as TNF,
interferon, KRAS or mTOR had either a minor and unreproduced association with iPFS

1
2
3
4 in only one cohort but more generally no significant correlation with survival in any of
5 the two ICI-treated MSI mCRC cohorts (**SUPPLEMENTARY FIGURE S5**).

6
7 In order to identify context-dependent transcriptomic signatures (*i.e.*, phenotypic
8 descriptors effectively observable in MSI mCRC), an unsupervised blind source
9 separation approach was applied to the transcriptome profiles of the discovery cohort
10 (C1; *i.e.*, independent component analysis; ICA) (**SUPPLEMENTARY FIGURE S6**).
11 Only one of these components was associated to shorter iPFS in both C1 and C2
12 ($HR_{C1} = 1.87$; 95% CI [1.06-3.3] ; $p = 0.0282$ - $HR_{C2} = 1.75$; 95% CI [1.03-2.98] ; $p =$
13 0.035) (**Fig. 3B and 3C and SUPPLEMENTARY FIGURE S6**). The component was
14 significantly correlated to an RNA-based quantification of fibroblasts ($p < 0.001$) and
15 anti-correlated to the tumor cellularity ($p < 0.001$) confirming its stromal origin (**FIG. 3A,**
16 **SUPPLEMENTARY TABLE S3** for the list of genes most contributing to this stromal
17 component). This fibrosis signature was associated to a tendon-like phenotype,
18 extracellular matrix (ECM)-producing and interacting genes, enriched intra-tumor
19 CMS4 proportions, and a pan-fibroblast TGF- β response signature (PMID: 29443960)
20 (**FIG. 3A and SUPPLEMENTARY FIGURE S7**). The fibrosis signature was not related
21 to stromal abundance, as measured by Sirius Red staining (**SUPPLEMENTARY**
22 **FIGURE S8**). Exactly as this was previously performed with the DNA signature, we
23 confirmed the independent prognostic value of this RNA signature using a multivariate
24 analysis in C2 ($HR_{C2} = 1.78$; 95% CI [1.03-3.01] ; $p = 0.040$) (**Fig. 3D and**
25 **SUPPLEMENTARY TABLE S1**). We successfully reduced this RNA signature to a
26 more serviceable set of 182 biomarkers (**SUPPLEMENTARY TABLE S4**). Using this
27 raw sub panel, the spearman correlation with the original component was 0.98 in
28 NIPICOL and 0.96 in immunoMSI (**SUPPLEMENTARY FIGURE S9**). Both remained
29 associated with iPFS as continuous scores. **SUPPLEMENTARY FIGURE S9** shows
30 the Kaplan-Meier curves of iPFS using the same cut-off as with the original signature.
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52 **Multivariate Analyses combining the DNA and RNA signatures**

53
54 We finally examined the effect of both RNA and DNA signatures in predicting iPFS
55 to ICI in MSI mCRC patients using a multivariate model (**FIG. 4**). Either when taking
56 into account both signatures in a continuous or a discrete way, the multivariate analysis
57 indicated that both signatures were independently predictive of iPFS (**FIG. 4A**). **Figures**
58
59
60

1
2
3
4 **4B** (and **SUPPLEMENTARY FIGURE S9**) show how these DNA and RNA signatures
5 are complementary and effective when used together in predicting the response of
6 patients to ICI treatment. One and/or another of the signatures was able to predict
7 progression in patients from C2 regarding iPFS to ICI ($HR_{C2} = 6.0$; 95% CI [1.80 –
8 19.9] $p = 9.12 \times 10^{-4}$). This makes these signatures a valuable classifier tool capable of
9 predicting relapse in MSI mCRC patients (see also on **SUPPLEMENTARY FIGURE**
10 **S10** the performance of this combined analysis when the RNA stromal signature was
11 reduced to 182 markers).
12
13
14
15
16

17 18 **DISCUSSION**

19
20
21
22
23 In this study, we examine the issue of response to ICI in patients with dMMR/MSI
24 mCRC. Two prospective cohorts of patients are examined to investigate this question.
25 The study is limited to CRC patients only, thus avoiding the risk that the origin of the
26 primary tumor might interfere with the results. The dual approach of high-throughput
27 analysis of tumor DNA and RNA enables us to perform both supervised and
28 unsupervised analyses in the aim of investigating ICI-response associated genotypes
29 and phenotypes. Note worthily, the dMMR/MSI status of the tumors was systematically
30 rechecked in our expert center as part of this ancillary study to avoid misdiagnoses
31 which can deeply impact the findings of such translational research studies, as
32 reported^{24, 35}. All these points constitute methodological strengths of the present work
33 compared to others in the field which remain few and based on smaller number of
34 patients.
35
36
37
38
39
40
41
42
43

44
45 We first show that several previously suggested DNA and RNA indicators are
46 unlikely to have any robust predictive values regarding response to ICI in MSI mCRC
47 patients. This is primarily the case for the quantitative DNA markers MSIsensor score
48 and TMB^{15, 16} when the diagnostic errors regarding the MSI status of the tumors are
49 corrected upstream and false positive MSI cases with expected low MSIsensor score
50 and TMB are therefore excluded. In addition, our analyses based on previously defined
51 candidate DNA/RNA markers of resistance, e.g., specific somatic DNA variants, the
52 activity of canonical signaling pathways or the presence of specific cellular contingents
53 within the tumor bulk, also failed to identify robust predictors of ICI response in patients.
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Though some of these markers may contribute to some extent to modulate the patients' response to ICI in some dMMR/MSI mCRC as suggested by some studies carried out on smaller series of patients, and more generally may be predictive of ICI efficacy in MMR proficient tumors, they are unlikely to have major clinical relevance.

The breakthrough results we have achieved in this study stem from the fact that we report and validate in both cohorts two distinct original signatures derived from a global analysis of the DNA and RNA profiles of ICI-treated MSI mCRC. In brief, the DNA signature is epithelial in nature. It analyzes the combination of a 19-plex panel aggregating independent somatic mutational events in dMMR tumor cells involving both coding and noncoding microsatellite containing-genes associated to distinct biological processes. As some microsatellite mutations such as those in the pentaplex are relevant for MSI diagnosis in CRC, the combined analysis of these 19 microsatellites is of predictive interest for response to ICI as evaluated by iPFS in patients. In addition, the RNA signature is stromal in nature. Though it is TGFB-related with a desmoplastic orientation, the fact that Sirius Red staining does not discriminate between samples with and without this stromal signature illustrates its complex nature. Note worthily, all the previously defined signatures and pathways that were shown to be associated to the proposed transcriptomic-defined TGFB-related fibrotic signature were not themselves associated to progression in both cohorts. This highlights the importance of defining context-dependent signatures, in particular in such distinct carcinogenic and microenvironment context as those found in MSI mCRC. In relation to a rich literature that has already provided evidence that the extracellular matrix and its heterogeneous content could promote resistance to ICI in MSS tumor models associated with various primary locations (for review see ³⁶), the RNA contingent we are identifying could thus play such a role by promoting, for example, immune-exclusion or sequestration processes in the context of increased fibrosis generated in and around the tumor. It could also modulate the response to treatment through non-mutually exclusive mechanisms involving more specifically the cancer-associated fibroblasts hosted in or responsible for the matrix generation.

Our work has some limitations. First, it does not address the issue of the choice of immunotherapy, i.e., monotherapy with anti-PD1 or combination of anti-PD1 and anti-CTL4, which is an important issue today. Second, the question of epigenetics is an

1
2
3 especially important issue that is not here addressed and will require further
4 investigation in subsequent studies. Third, our work does not really shed light on simple
5 mechanisms underlying resistance to ICI in MSI mCRC. In this respect, it is not yet
6 very enlightening for the identification of novel therapeutic targets that could be of
7 interest in the future for patients with primary or secondary resistance to anti-PD-1 ±
8 anti-CTLA4. Finally, although our results were established on a large prospective
9 collection of patients including two independent cohorts, one of which is a clinical trial,
10 they still require validation in new populations of metastatic MSI cancer patients, with
11 colorectal tumor but possibly also other primary MSI cancer. Despite these
12 weaknesses, the 2 signatures we outline here from the analysis of the tumor bulk are
13 easy to investigate and can be implemented in the context of a clinical routine; the
14 DNA signature requires the analysis of the status of only 19 microsatellite markers
15 within the tumor, ideally by NGS but possibly also by other methods while the RNA
16 signature requires to perform 3'RNA-sequencing which is a simple, reproducible and
17 feasible method to be used with paraffin-embedded tumor samples, the cost of which
18 has now decreased considerably. Of interest, we demonstrate that the RNA signature
19 reduced to 182 markers is still clinically relevant for predicting iPFS to ICI. Finally, it is
20 very interesting to note that the DNA and RNA signatures have independent predictive
21 value and can be used jointly to predict progression in ICI-treated MSI mCRC patients.
22 This should be of interest for the design of future strategies improving ICI for MSI
23 cancer patients in clinics.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

46 REFERENCES

- 47 1 Ionov Y, Peinado MA, Malkhosyan S et al. Ubiquitous somatic mutations in
48 simple repeated sequences reveal a new mechanism for colonic
49 carcinogenesis. *Nature* 1993; 363 (6429): 558-561.
- 50 2 Aaltonen LA, Peltomaki P, Leach FS et al. Clues to the pathogenesis of familial
51 colorectal cancer. *Science* 1993; 260 (5109): 812-816.
- 52 3 Fishel R, Lescoe MK, Rao MR et al. The human mutator gene homolog MSH2
53 and its association with hereditary nonpolyposis colon cancer. *Cell* 1993; 75 (5):
54 1027-1038.
55
56
57
58
59
60

- 1
- 2
- 3
- 4 Svrcek M, Lascols O, Cohen R et al. MSI/MMR-deficient tumor diagnosis: Which standard for screening and for diagnosis? Diagnostic modalities for the colon and other sites: Differences between tumors. *Bull Cancer* 2019; 106 (2): 119-128.
- 5 Hamelin R, Chalastanis A, Colas C et al. [Clinical and molecular consequences of microsatellite instability in human cancers]. *Bull Cancer* 2008; 95 (1): 121-132.
- 6 Duval A, Hamelin R. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. *Cancer Res* 2002; 62 (9): 2447-2454.
- 7 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; 5 (1): 43-51.
- 8 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.
- 9 Cohen R, Colle R, Pudlarz T et al. Immune Checkpoint Inhibition in Metastatic Colorectal Cancer Harboring Microsatellite Instability or Mismatch Repair Deficiency. *Cancers (Basel)* 2021; 13 (5).
- 10 Andre T, Shiu KK, Kim TW et al. Pembrolizumab in Microsatellite-Instability-High Advanced Colorectal Cancer. *N Engl J Med* 2020; 383 (23): 2207-2218.
- 11 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.
- 12 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.
- 13 Overman MJ, Lonardi S, Wong KYM et al. Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair-Deficient/Microsatellite Instability-High Metastatic Colorectal Cancer. *J Clin Oncol* 2018; 36 (8): 773-779.
- 14 Cohen R, Bennouna J, Meurisse A et al. RECIST and iRECIST criteria for the evaluation of nivolumab plus ipilimumab in patients with microsatellite

- 1
2
3 instability-high/mismatch repair-deficient metastatic colorectal cancer: the
4 GERCOR NIPICOL phase II study. *J Immunother Cancer* 2020; 8 (2).
5
6
7 15 Mandal R, Samstein RM, Lee KW et al. Genetic diversity of tumors with
8 mismatch repair deficiency influences anti-PD-1 immunotherapy response.
9 *Science* 2019; 364 (6439): 485-491.
10
11
12 16 Schrock AB, Ouyang C, Sandhu J et al. Tumor mutational burden is predictive
13 of response to immune checkpoint inhibitors in MSI-high metastatic colorectal
14 cancer. *Ann Oncol* 2019; 30 (7): 1096-1103.
15
16
17 17 Gurjao C, Liu D, Hofree M et al. Intrinsic Resistance to Immune Checkpoint
18 Blockade in a Mismatch Repair-Deficient Colorectal Cancer. *Cancer Immunol*
19 *Res* 2019; 7 (8): 1230-1236.
20
21
22 18 Chida K, Kawazoe A, Suzuki T et al. Transcriptomic Profiling of MSI-H/dMMR
23 Gastrointestinal Tumors to Identify Determinants of Responsiveness to Anti-
24 PD-1 Therapy. *Clin Cancer Res* 2022; 28 (10): 2110-2117.
25
26
27 19 Kwon M, An M, Klempner SJ et al. Determinants of Response and Intrinsic
28 Resistance to PD-1 Blockade in Microsatellite Instability-High Gastric Cancer.
29 *Cancer Discov* 2021; 11 (9): 2168-2185.
30
31
32 20 Bortolomeazzi M, Keddar MR, Montorsi L et al. Immunogenomics of Colorectal
33 Cancer Response to Checkpoint Blockade: Analysis of the KEYNOTE 177 Trial
34 and Validation Cohorts. *Gastroenterology* 2021; 161 (4): 1179-1193.
35
36
37 21 Chida K, Kawazoe A, Kawazu M et al. A Low Tumor Mutational Burden and
38 PTEN Mutations Are Predictors of a Negative Response to PD-1 Blockade in
39 MSI-H/dMMR Gastrointestinal Tumors. *Clin Cancer Res* 2021; 27 (13): 3714-
40 3724.
41
42
43 22 Latham A, Srinivasan P, Kemel Y et al. Microsatellite Instability Is Associated
44 With the Presence of Lynch Syndrome Pan-Cancer. *J Clin Oncol* 2019; 37 (4):
45 286-295.
46
47
48 23 Germano G, Lu S, Rospo G et al. CD4 T Cell-Dependent Rejection of Beta-2
49 Microglobulin Null Mismatch Repair-Deficient Tumors. *Cancer Discov* 2021; 11
50 (7): 1844-1859.
51
52
53 24 Cohen R, Hain E, Buhard O et al. Association of Primary Resistance to Immune
54 Checkpoint Inhibitors in Metastatic Colorectal Cancer With Misdiagnosis of
55
56
57
58
59
60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Microsatellite Instability or Mismatch Repair Deficiency Status. *JAMA Oncol* 2019; 5 (4): 551-555.
- 25 Colle R, Radzik A, Cohen R et al. Pseudoprogession in patients treated with immune checkpoint inhibitors for microsatellite instability-high/mismatch repair-deficient metastatic colorectal cancer. *Eur J Cancer* 2021; 144: 9-16.
- 26 Buhard O, Cattaneo F, Wong YF et al. Multipopulation analysis of polymorphisms in five mononucleotide repeats used to determine the microsatellite instability status of human tumors. *J Clin Oncol* 2006; 24 (2): 241-251.
- 27 Buhard O, Lagrange A, Guilloux A et al. HSP110 T17 simplifies and improves the microsatellite instability testing in patients with colorectal cancer. *J Med Genet* 2016; 53 (6): 377-384.
- 28 Ratovomanana T, Cohen R, Svrcek M et al. Performance of Next-Generation Sequencing for the Detection of Microsatellite Instability in Colorectal Cancer With Deficient DNA Mismatch Repair. *Gastroenterology* 2021; 161 (3): 814-826 e817.
- 29 Seymour L, Bogaerts J, Perrone A et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol* 2017; 18 (3): e143-e152.
- 30 Boegel S, Lower M, Schafer M et al. HLA typing from RNA-Seq sequence reads. *Genome Med* 2012; 4 (12): 102.
- 31 Hundal J, Kiwala S, McMichael J et al. pVACtools: A Computational Toolkit to Identify and Visualize Cancer Neoantigens. *Cancer Immunol Res* 2020; 8 (3): 409-420.
- 32 Lundegaard C, Lamberth K, Harndahl M et al. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. *Nucleic Acids Res* 2008; 36 (Web Server issue): W509-512.
- 33 Reynisson B, Alvarez B, Paul S et al. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* 2020; 48 (W1): W449-W454.

- 1
2
3
4 34 Abraham A, Pedregosa F, Eickenberg M et al. Machine learning for
5 neuroimaging with scikit-learn. *Front Neuroinform* 2014; 8: 14.
6
7 35 Loupakis F, Depetris I, Biondo P et al. Prediction of Benefit from Checkpoint
8 Inhibitors in Mismatch Repair Deficient Metastatic Colorectal Cancer: Role of
9 Tumor Infiltrating Lymphocytes. *Oncologist* 2020; 25 (6): 481-487.
10
11
12 36 O'Donnell JS, Teng MWL, Smyth MJ. Cancer immunoediting and resistance to
13 T cell-based immunotherapy. *Nat Rev Clin Oncol* 2019; 16 (3): 151-167.
14
15
16 37 Powles T, Kockx M, Rodriguez-Vida A et al. Clinical efficacy and biomarker
17 analysis of neoadjuvant atezolizumab in operable urothelial carcinoma in the
18 ABACUS trial. *Nat Med* 2019; 25 (11): 1706-1714.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

FIGURE LEGENDS

Table 1: Patients description from C1 and C2 cohorts with main corresponding clinical data

Figure 1: Workflow and Study Design

MSS, microsatellite stable; mCRC, metastatic colorectal cancer; ICI, immune checkpoint inhibitors; IHC, Immunohistochemistry; QC, quality control; pMMR, mismatch repair proficient.

Figure 2: Identification of a 19-plex MSI signature predictive of response to ICI in mCRC patients

A) Histogram of p -values from cox analysis for R variants (microsatellites) in the C1 Nipicol cohort. A zoom was performed on those having a p -value lower than 0.1 ($n = 167$). Genes' name of the 19 somatic MS variants with a known percentage of missing data $< 5\%$ in C1 and other public and private cohorts are indicated in green. They were used further for investigating iPS to ICI in C2.

B) Counts of microsatellite somatic variants according to their p -value (uni-variate Cox regression model) on the 23 patients of Nipicol (C1). Green and red bar represent the mutated microsatellites with a p -value < 0.05 and $0.05 \leq p$ -value < 0.1 , respectively. The top forestplot represent in C2 the Cox survival regression results on model prediction of risk output, represented by the hazard ratio (HR < 1 : beneficial, HR = 1 No effect, HR > 1 : deleterious). High (p -value < 0.05 , DNA signature) in green vs low stringency of Cox features selection based on p -value in red.

C) iPFS curve for the TMB and the MSIsensore score.

D) Kaplan-Meier curves of iPFS are shown accordingly to risk probability in ImmunoMSI only ($n= 66$). Red curve corresponds to patients with high MNB probability ($<20\%$) and blue curve corresponds to patients with low MNB probability ($>20\%$).

E) Heatmap representation of NIPICOL (C1) and ImmunoMSI (C2). Overview of mutations profile for each patient on the list of 19 selected microsatellites (MS, see

genes' name and chromosomal location on the left). Bottom bars represent their treatment (combo-therapy or mono-therapy), survival (time-to-event), event (Progressor or Non-Progressor), DNA.signature (probability of non-response according to iPFS given by the MNB), KRAS and BRAF mutational status. Patients are ordered based on predicted risk of progress and MS are ordered in function of mutation frequency in Nipicol cohort.

F) Forest plot of the multivariate Cox regression with iPFS in cohort C2, with its corresponding p -value. Included features are DNA signature (low to high risk), BRAF, KRAS (wild-type vs mutated status) and treatment (mono vs. combo).

Figure 3: Transcriptomic signatures and response to ICI in MSI mCRC patients

A) Heatmap of the stromal signature, patients (in column) are ordered by the estimated quantification of the stromal ICA signature. iRECIST status of patients²⁹ are shown along with gene-wise centered expression values of the gene most associated to the component, tumor purity estimated by the non-repeated variant allele frequency, MCPcounter -based fibroblast quantification estimate, the main CMS (consensus molecular subtype) and the intra-tumoral proportion estimate of CMS4, and TGFB response signature³⁷. Pathway-level activity estimate are also shown along with the Gene Set Enrichment Analysis estimates.

B) Univariate association with progression-free survival in Nipicol (C1, n=44) or ImmunoMSI (C2, n=71) of the continuous score of the signature or of a classification using the third quartile as a threshold (Q3 cut-off).

C) Kaplan-Meier estimates using the third quantile threshold of the stromal signature.

D) Forest plot of the multivariate Cox regression with progression-free survival in cohort C2, representing hazard ratio (HR < 1 : beneficial, HR = 1 No effect, HR > 1 : deleterious) with its corresponding p .value. Included features are stromal signature (low to high), BRAF, KRAS (wild-type vs mutated), and treatment (mono vs. combo).

Figure 4: Multivariate Analyses combining the DNA and RNA signatures

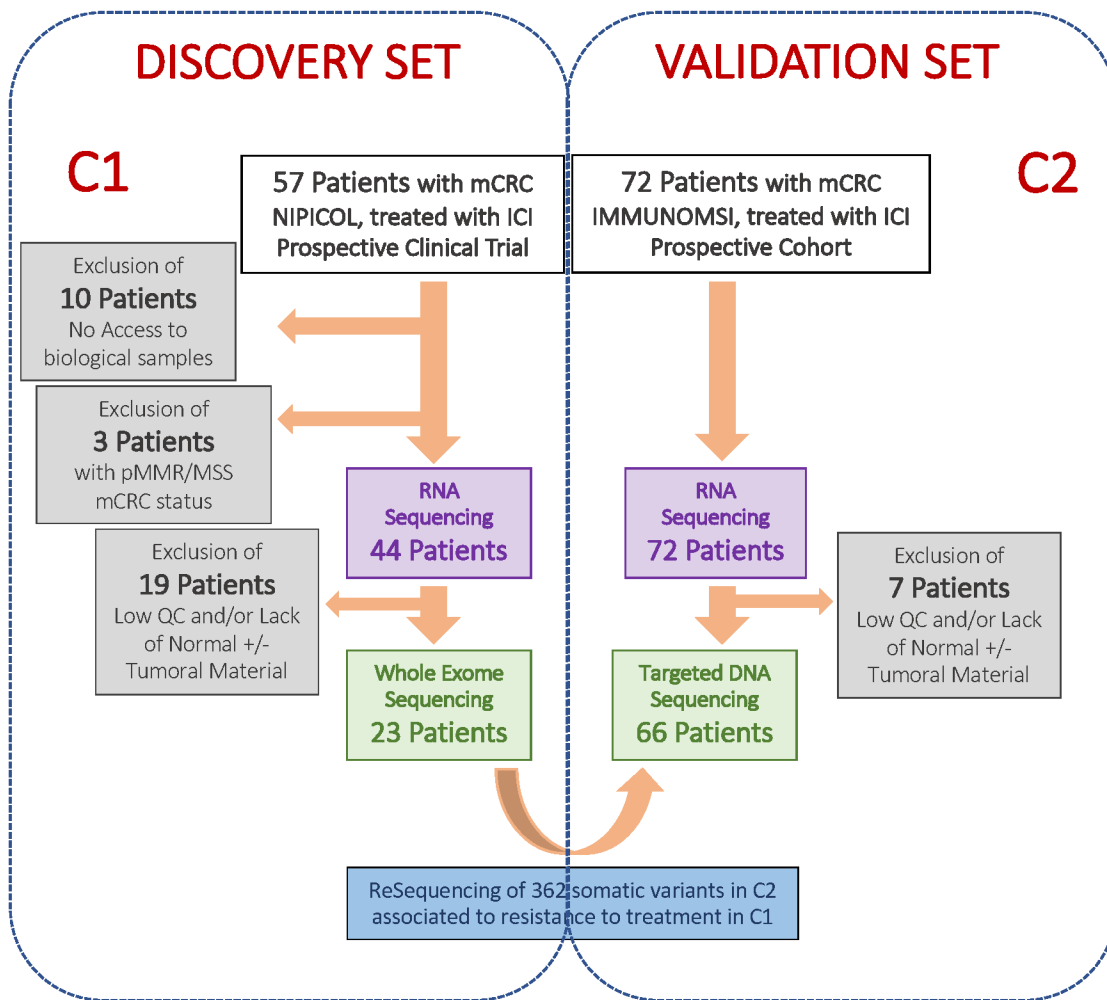
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

A) Forestplot of the RNA or DNA signatures as well as the treatment type considering the continuous value (upper panel) or by discretizing the values (lower panel) of the signatures.

B) Kaplan-Meier estimates of iPFS as a function of the intensity combination of the two signatures (n=66). High RNA and DNA signatures are shown in red. Low RNA and DNA signatures are shown in green. High RNA signature and weak DNA signature are shown in orange and low RNA signature and high DNA signature are shown in purple.

For Peer Review

A. Workflow



B. Study Design

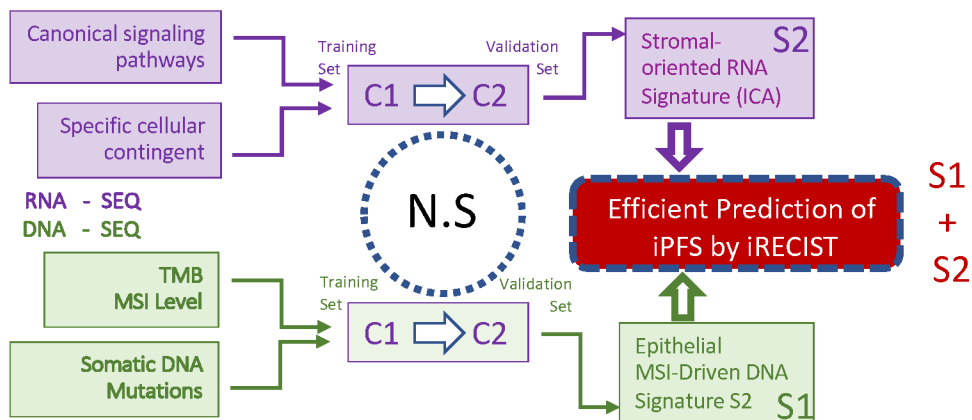
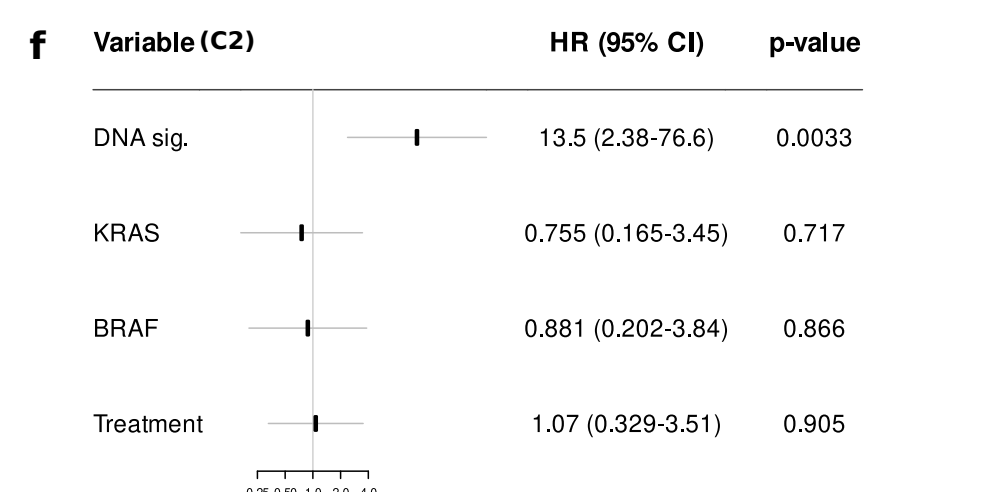
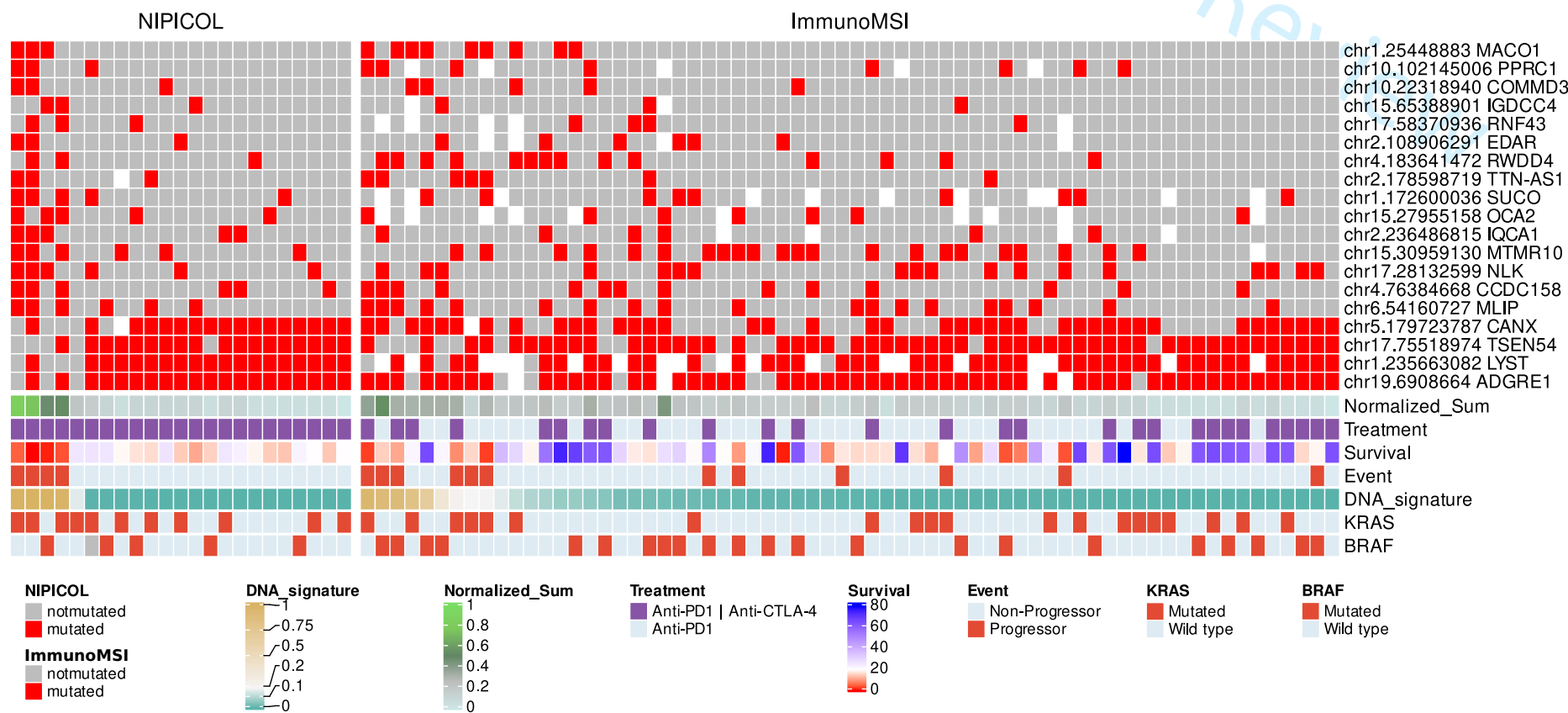
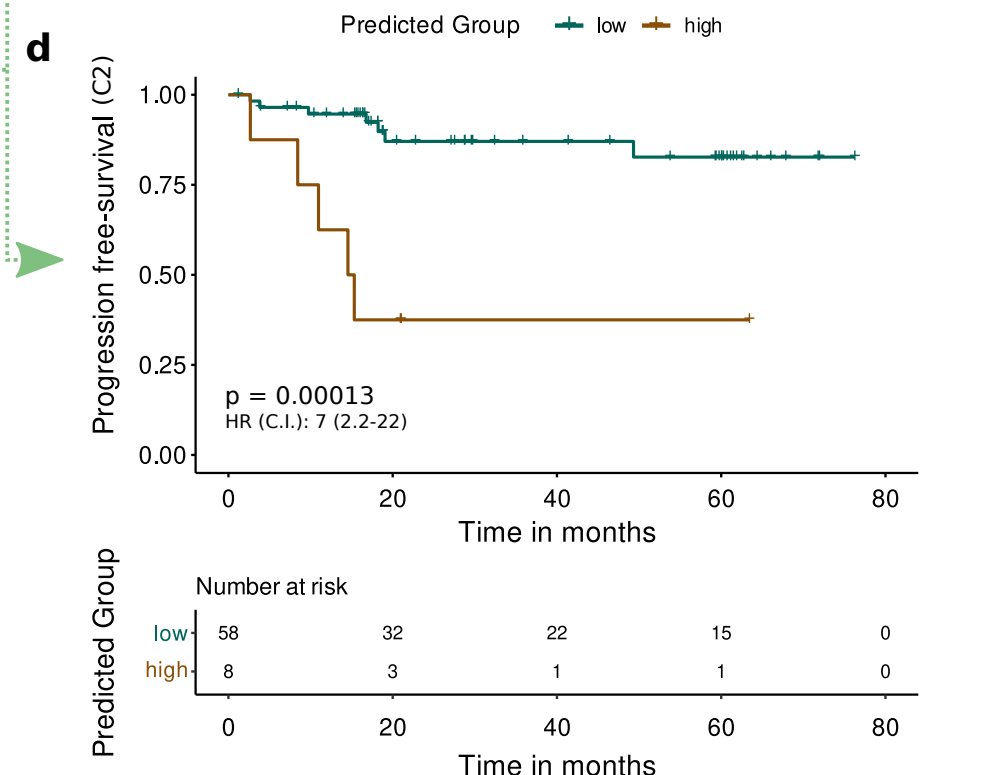
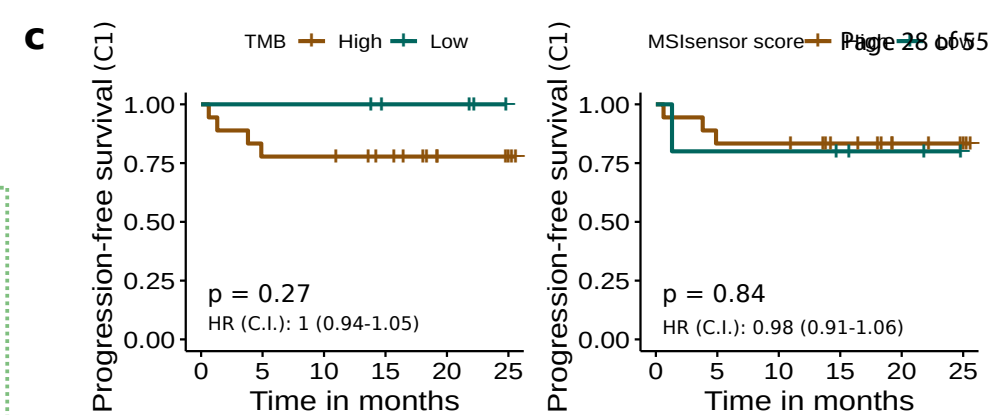
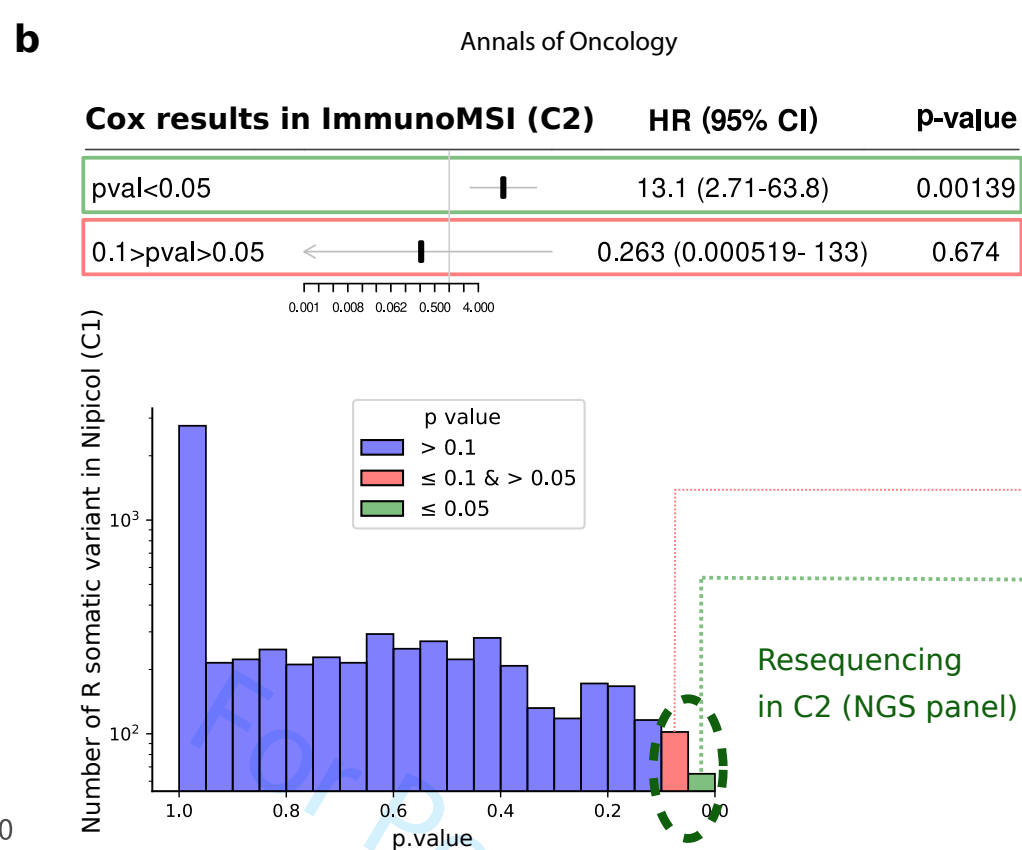
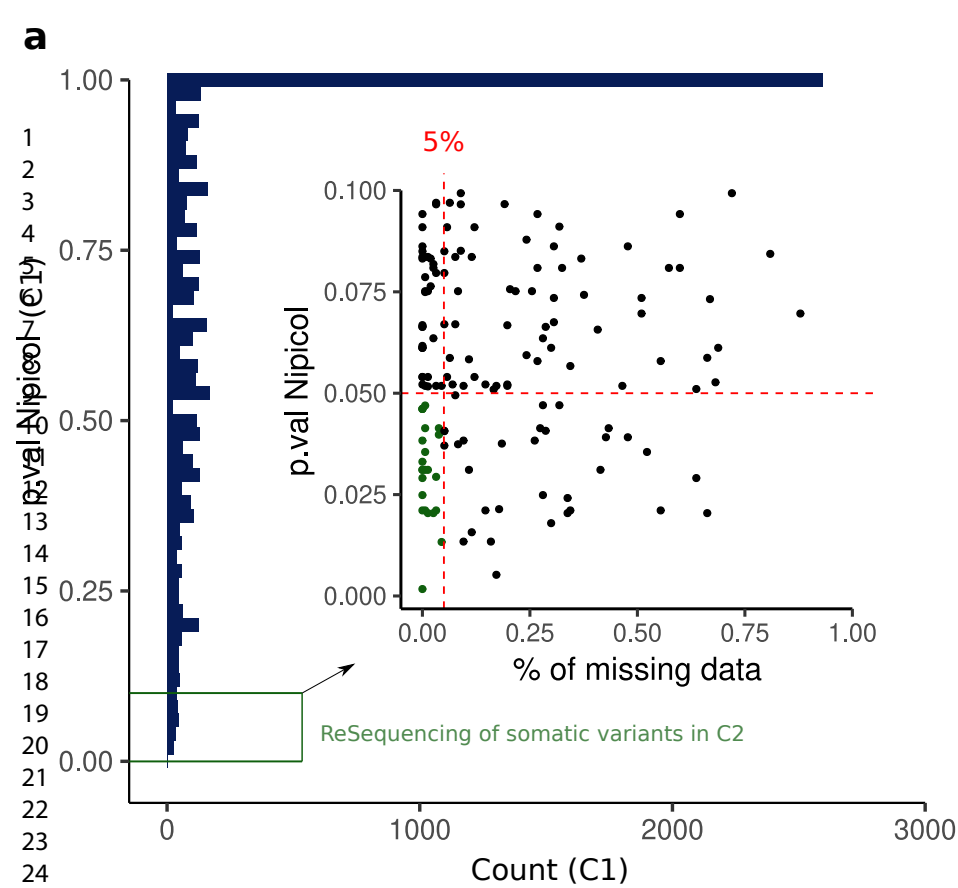


Figure 1



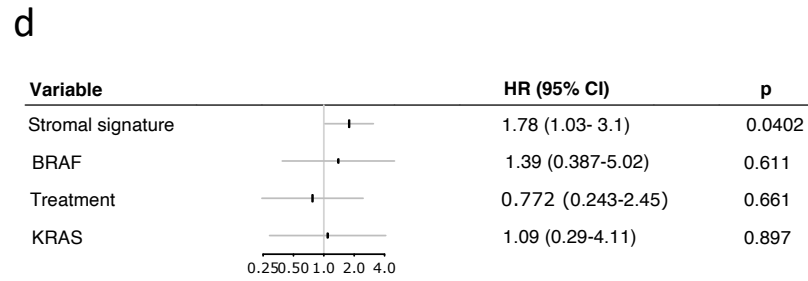
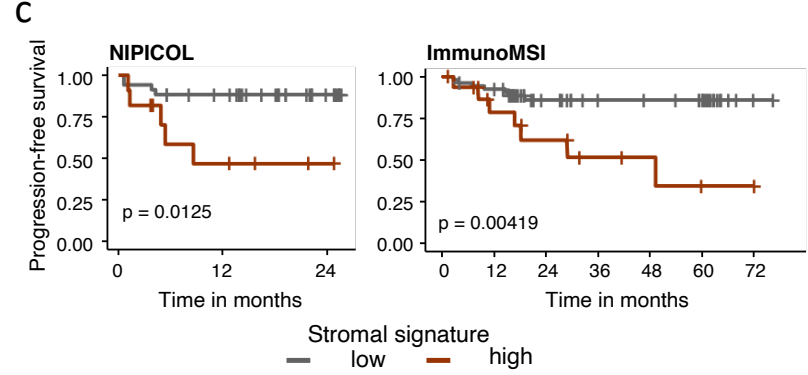
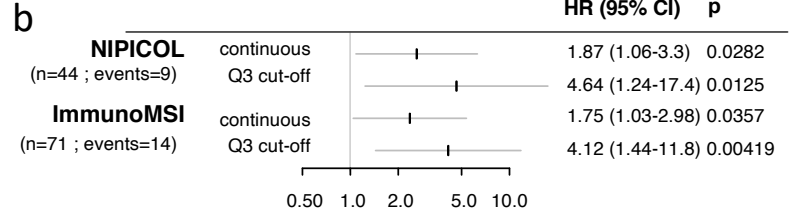
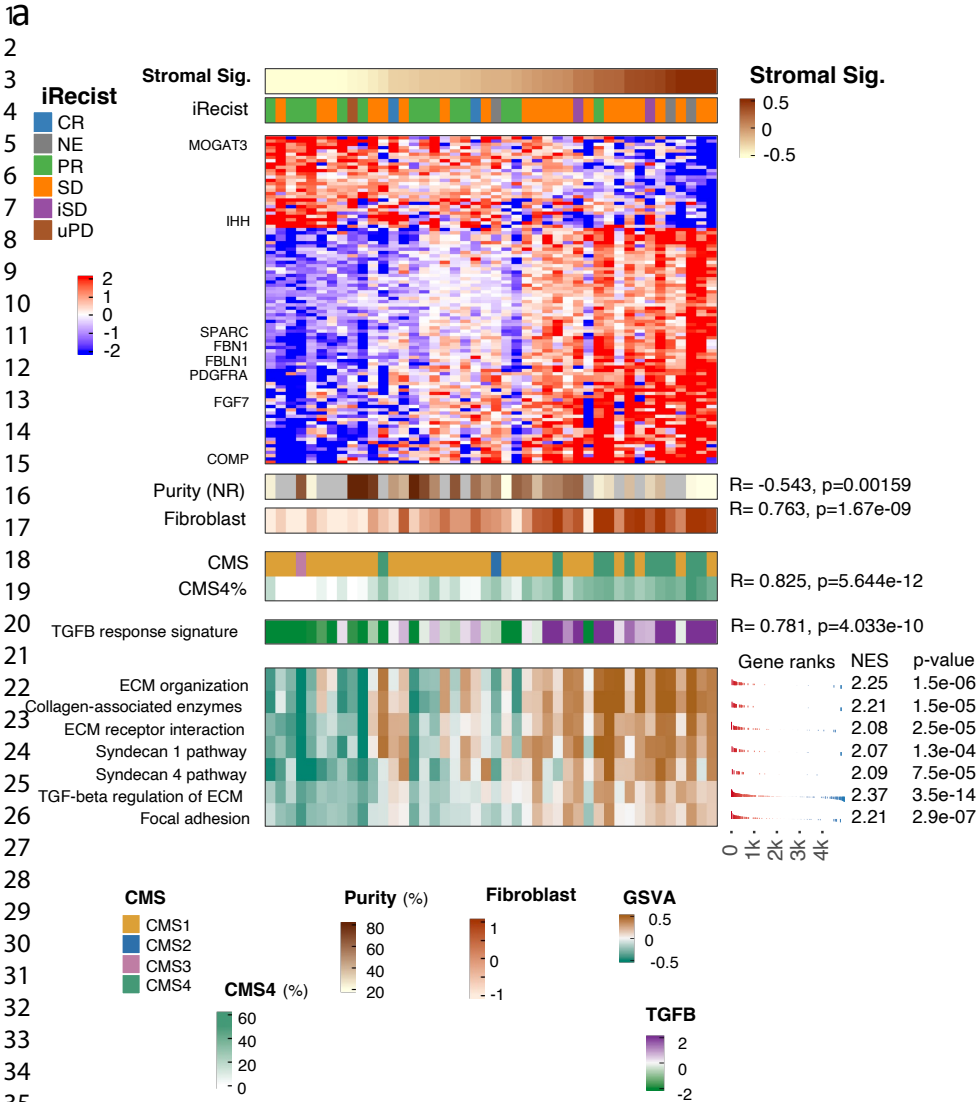
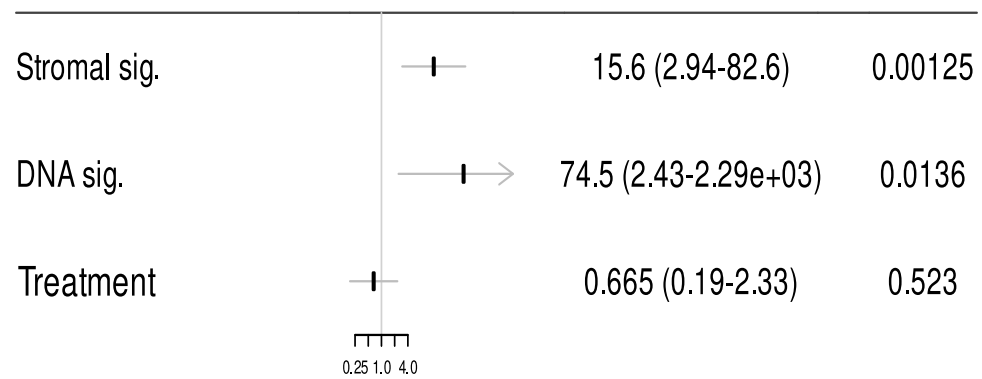
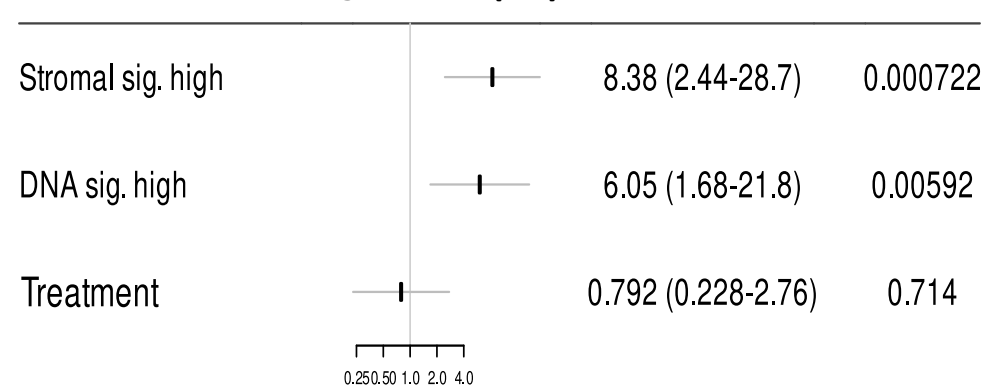
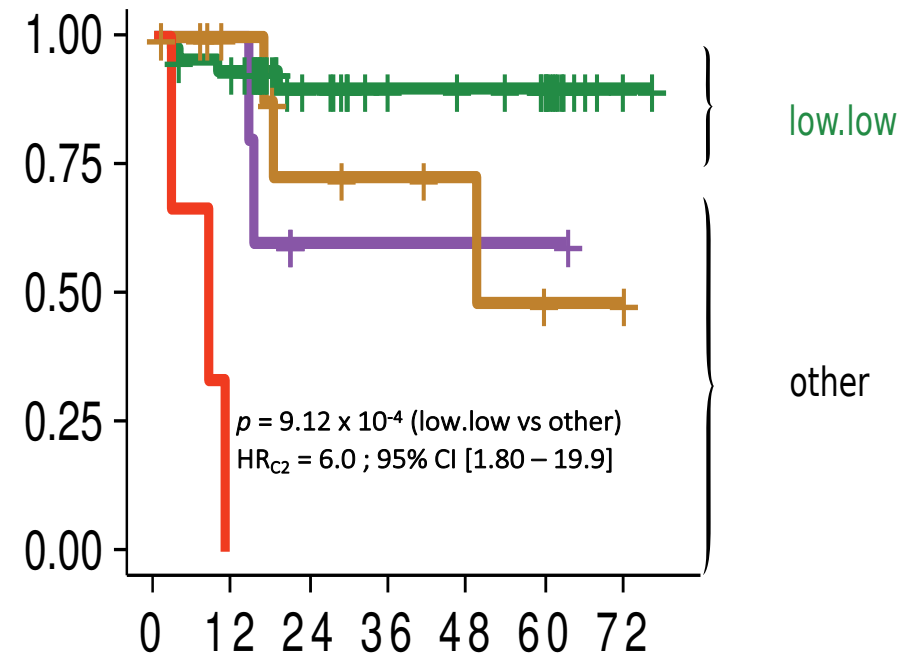


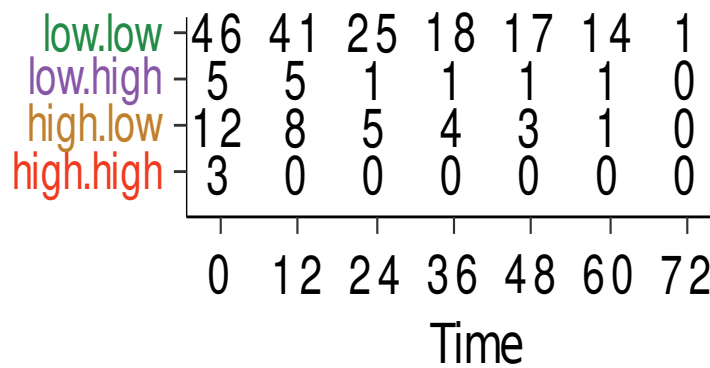
Figure 3

a**Continuous signatures (C2)
in multivariate****Multivariatecutoff signatures (C2)****b**

Progression-free survival (c2)



RNA.DNA

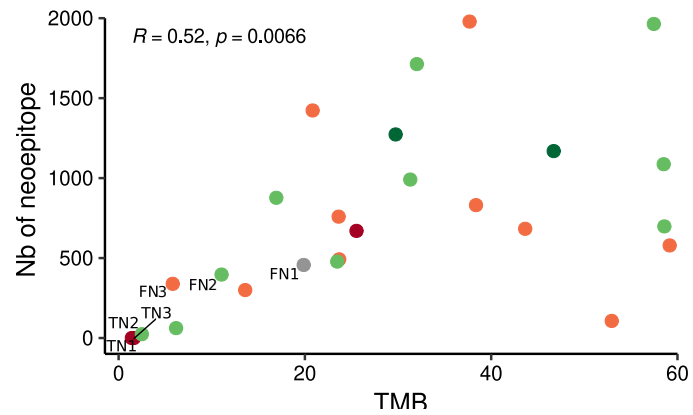
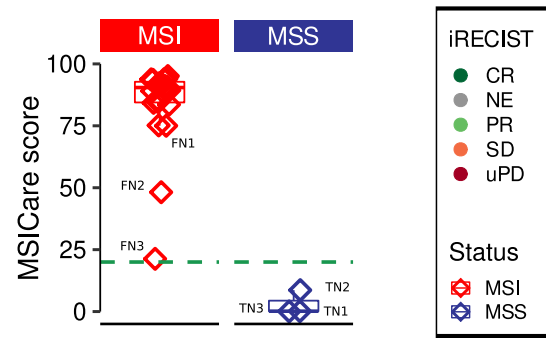
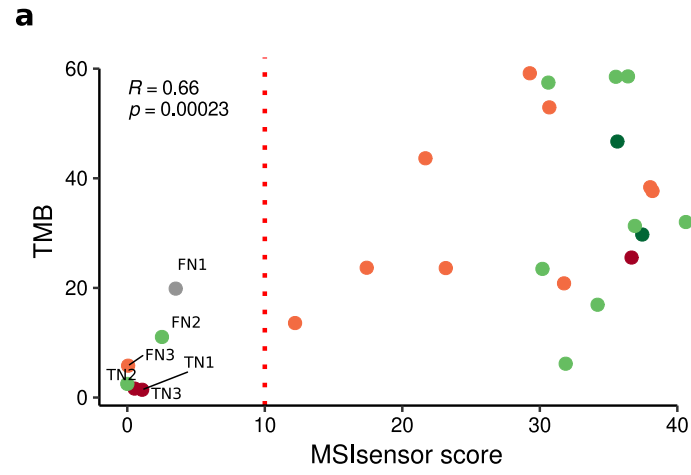
**Figure 4**

Characteristic	ImmunoMSI, N = 72 ¹	Nipicol, N = 47 ¹	p- value ²
Age in Yrs	59 (15)	54 (13)	0.058
Sex			0.87
Female/Male	28/44	19/28	
ECOG performance score			0.24
0/≥1/Unknown	35/36/1	18/29/0	
Primary tumor location			0.49
Left colon/ Right colon/Unknown	23/47/2	18/28/1	
BRAFV600E			0.056
Mutated/ Wild type/Unknown	23/49/0	7/37/3	
KRAS			0.029
Mutated/ Wild type	21/51	23/24	
NRAS			>0.99
Mutated/Wild type/Unknown	1/71/0	1/42/4	
Origin of MMR deficiency			0.82
Known germline mutation	20	27	
Sporadic/Unknown	10/42	12/8	
Surgery of primary tumor			>0.99
Yes/no/Unknown	71/1/0	43/1/3	
Number of metastatic sites			0.50
1 /2/ ≥3	27/26/19	14/18/15	
Number of prior lines			0.006
0/1/2/≥3	9/21/32/10	0/8/22/17	
Type of immunotherapy			<0.001
Anti-PD1/anti-PD1&CTLA4	43/29	0 / 47	
iRECIST			<0.001
CR/PR/SD/PD/NE	20/35/7/8/2	2/15/24/3/3	
Status			0.059
MSI/MSS	72 / 72 (100%)	44 / 3	

¹ Mean (SD); n / N (%)

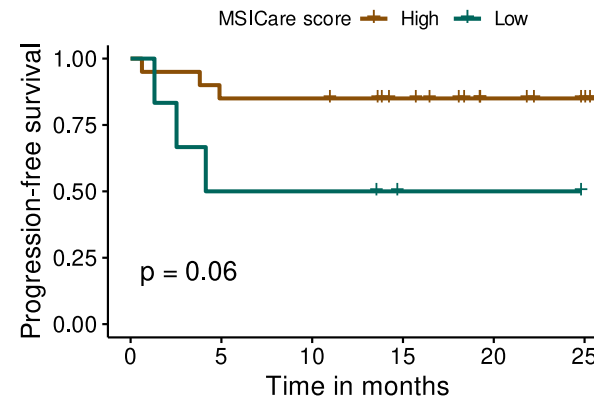
² Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

Yrs: years; ECOG = Eastern Oncology Cooperative Group; MMR: Mismatch Repair; immune RECIST: iRECIST CR: complete response; PR: partial response; SD: stable disease; PD: progressive disease; MSI: microsatellite instability; MSS: microsatellite stable, NE: not evaluable

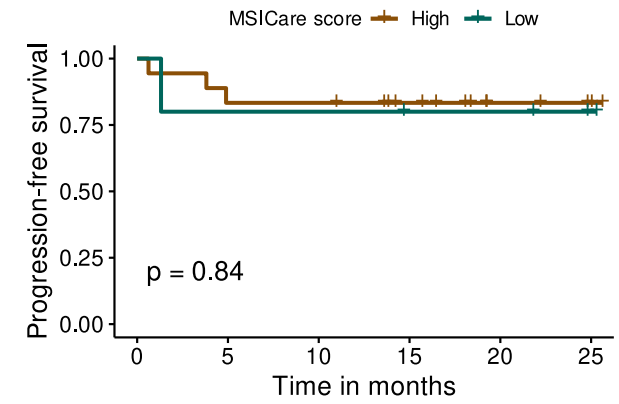


b

Before exclusion of MSS samples



After exclusion of MSS samples



Before exclusion of MSS samples

MSIsensor_score

HR (95% CI)

0.963 (0.914-1.02)

p-value

0.165

TMB

0.975 (0.93-1.02)

0.282

Neoepitope

0.999 (0.997- 1)

0.3

0.71 1.0

After exclusion of MSS samples

MSIsensor_score

HR (95% CI)

0.984 (0.912-1.06)

p-value

0.672

Neoepitope

1 (0.998- 1)

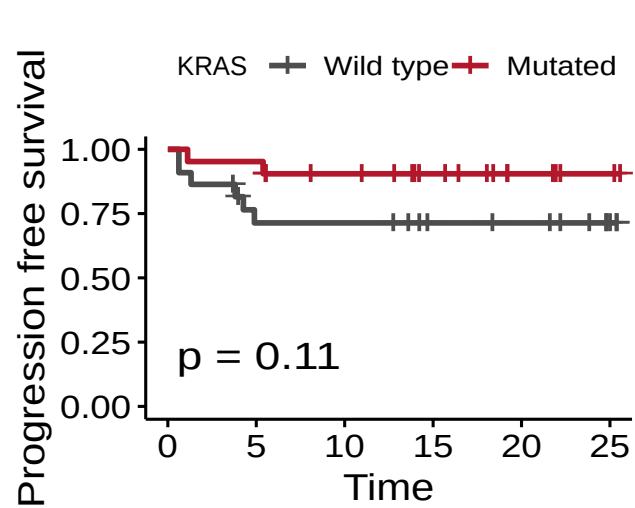
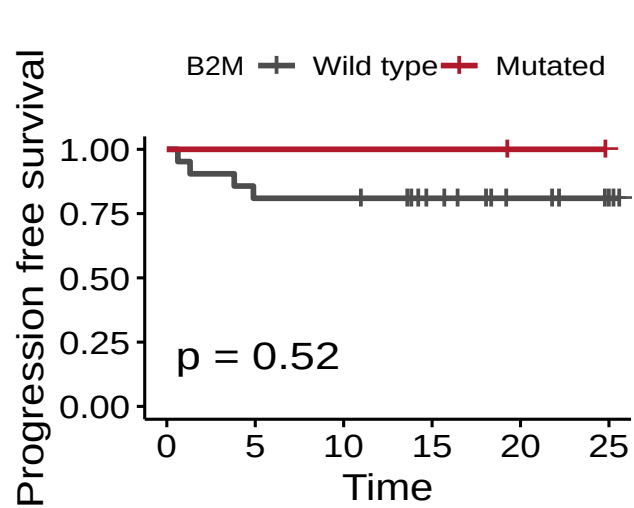
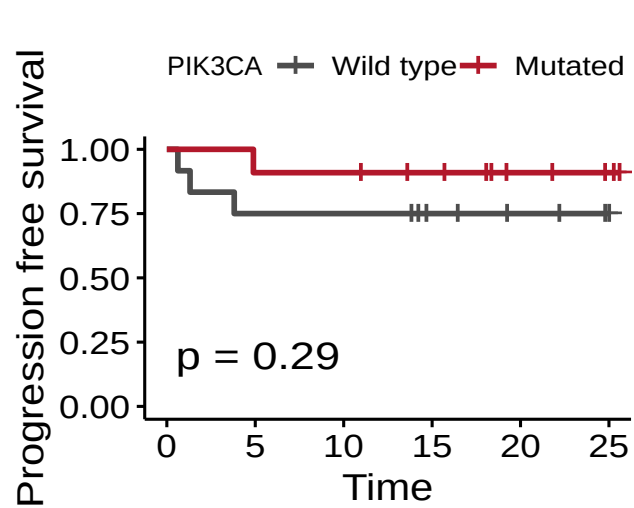
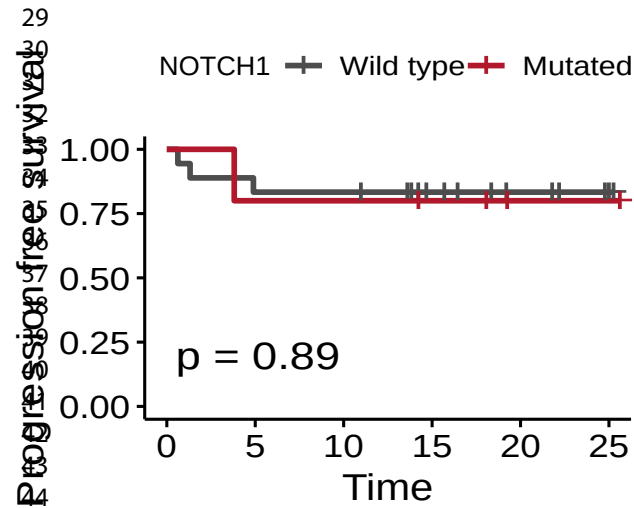
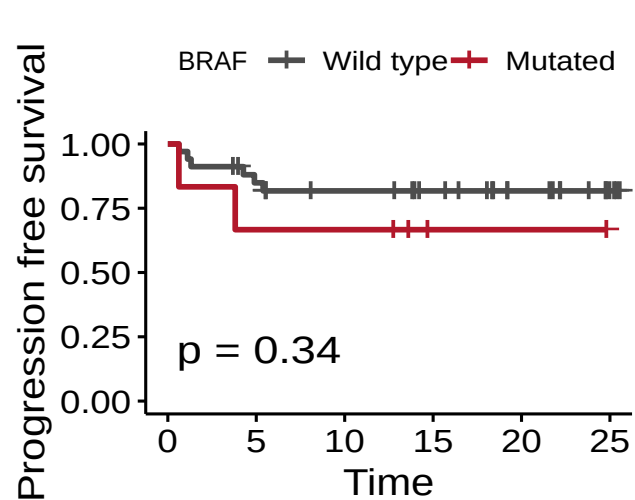
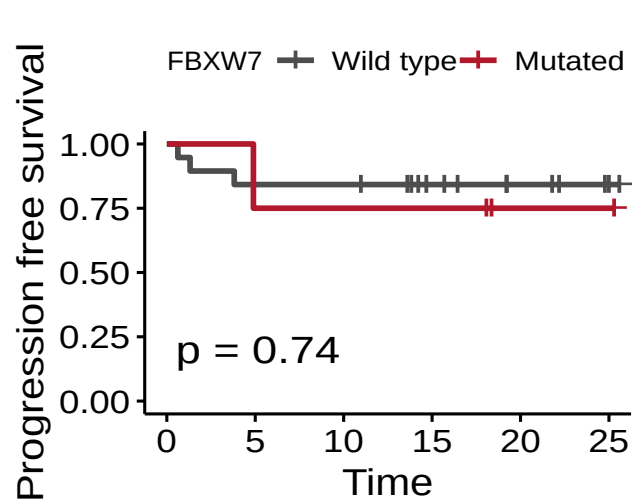
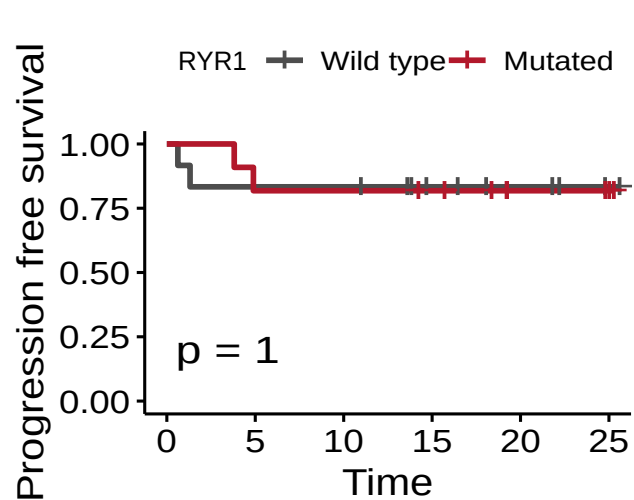
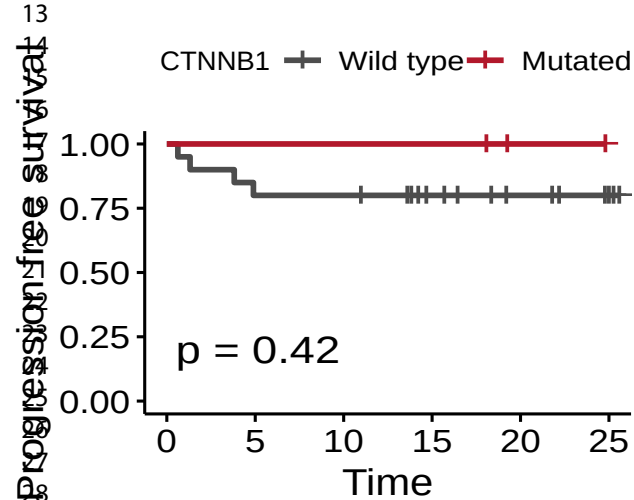
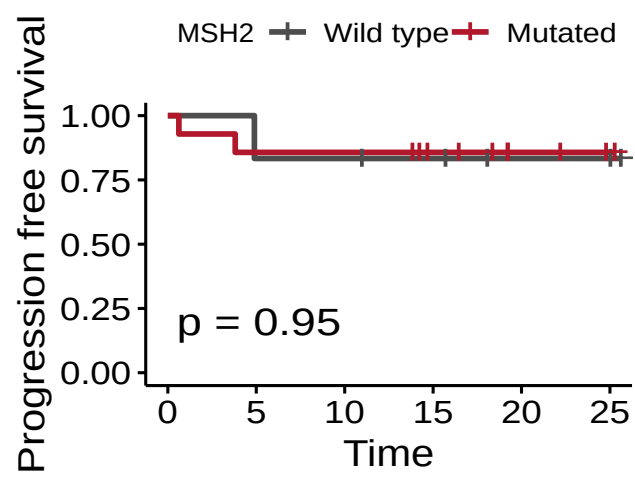
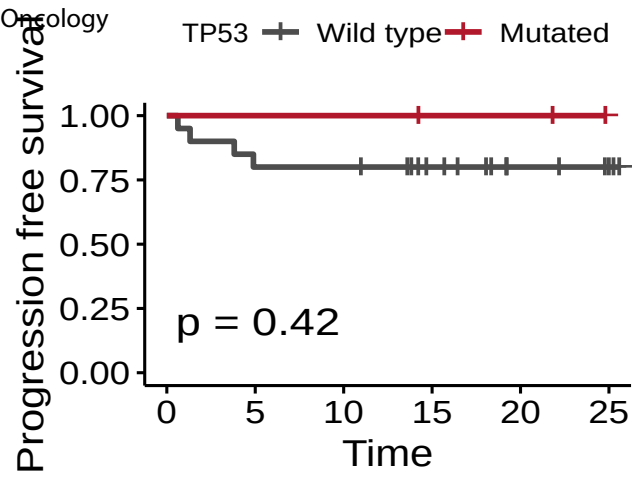
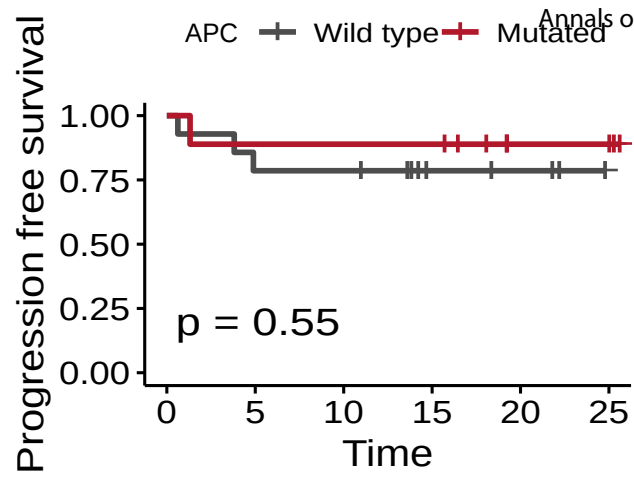
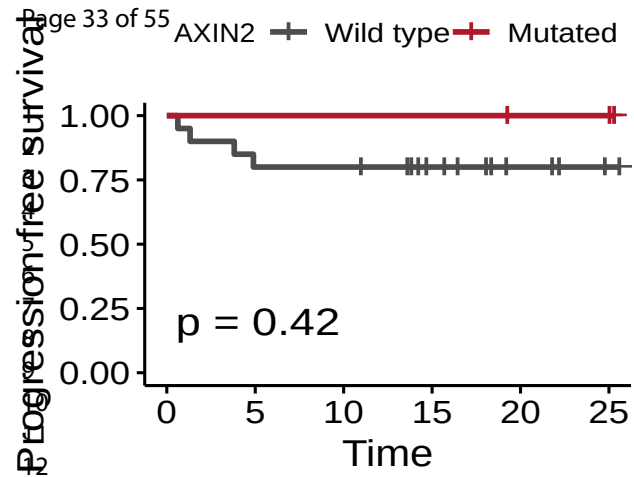
0.807

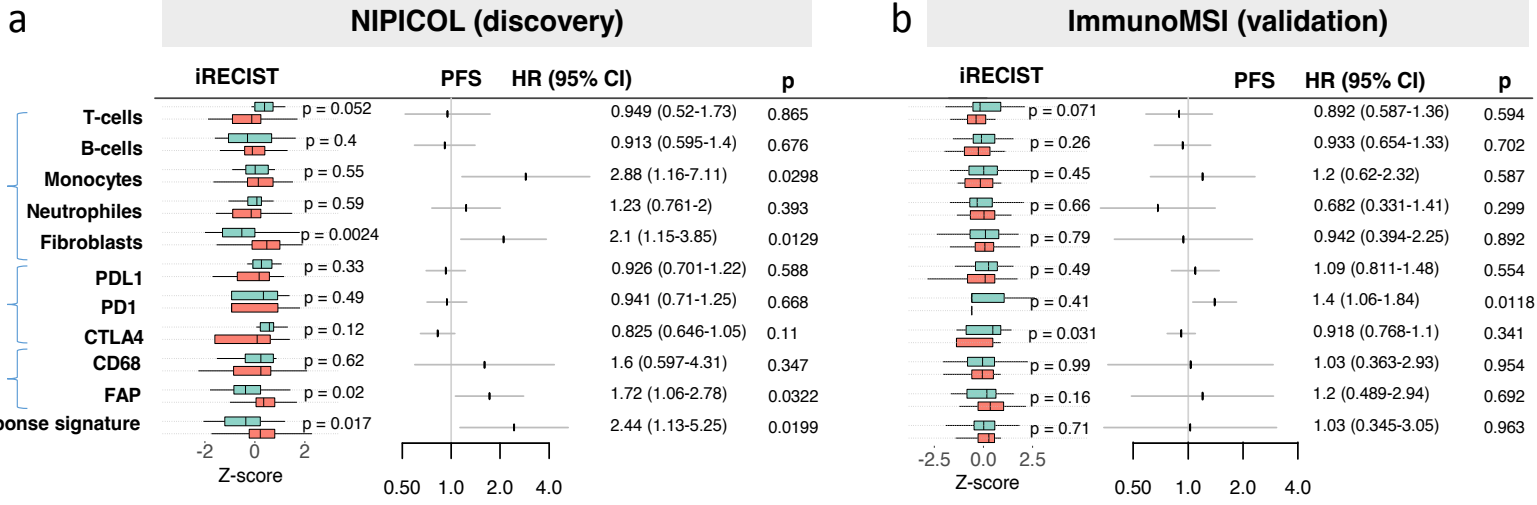
TMB

0.995 (0.94-1.05)

0.876

0.71 1.0

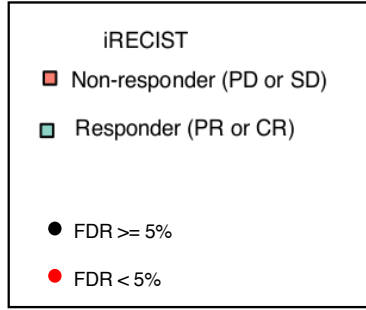
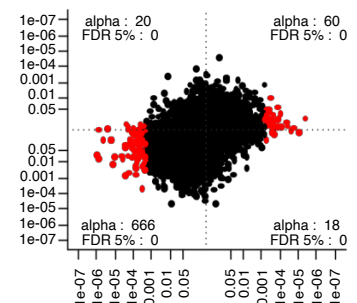
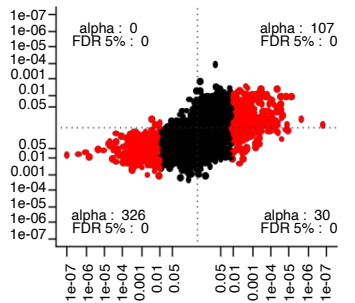
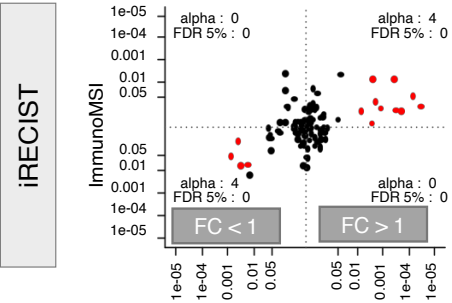




C TME

d Pathway

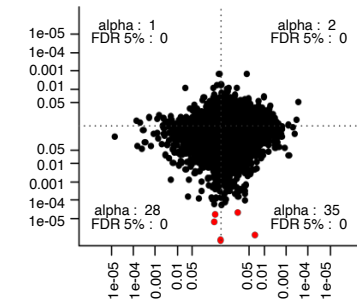
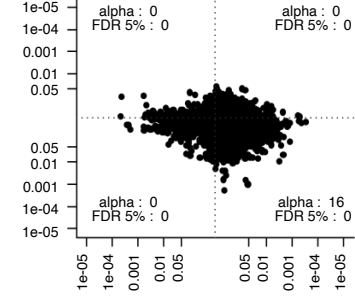
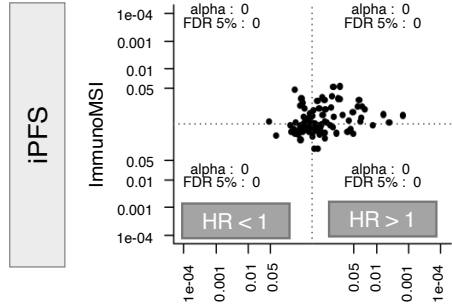
e Genes



f

g

h



NIPICOL

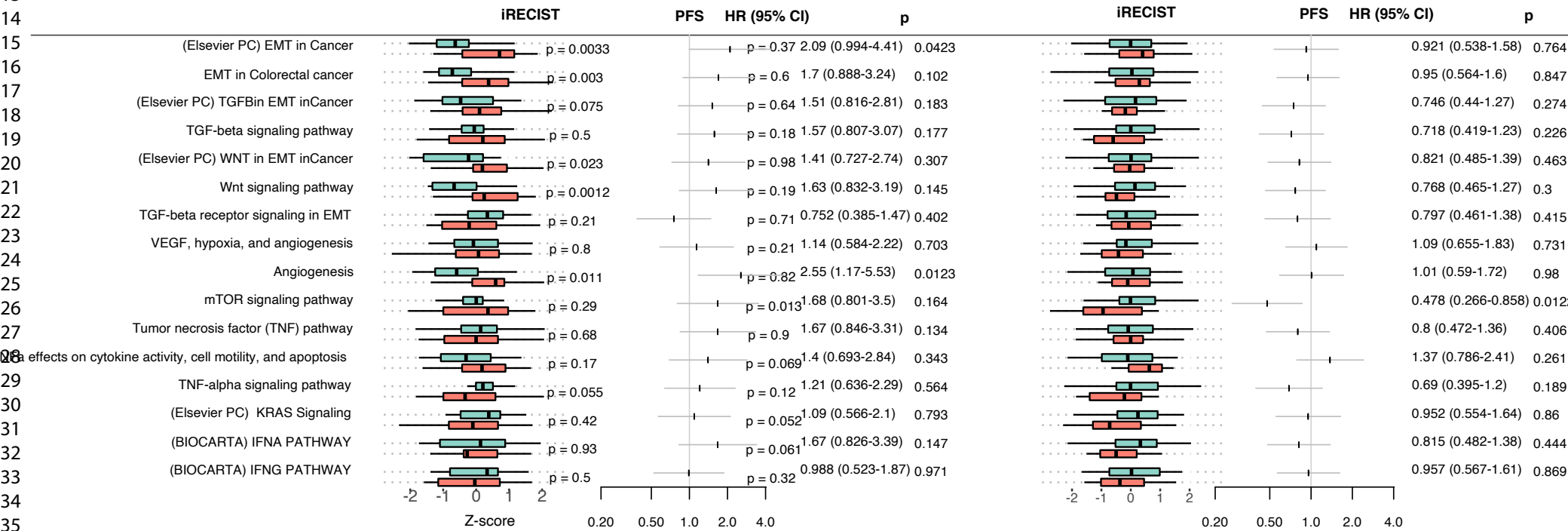
NIPICOL

NIPICOL

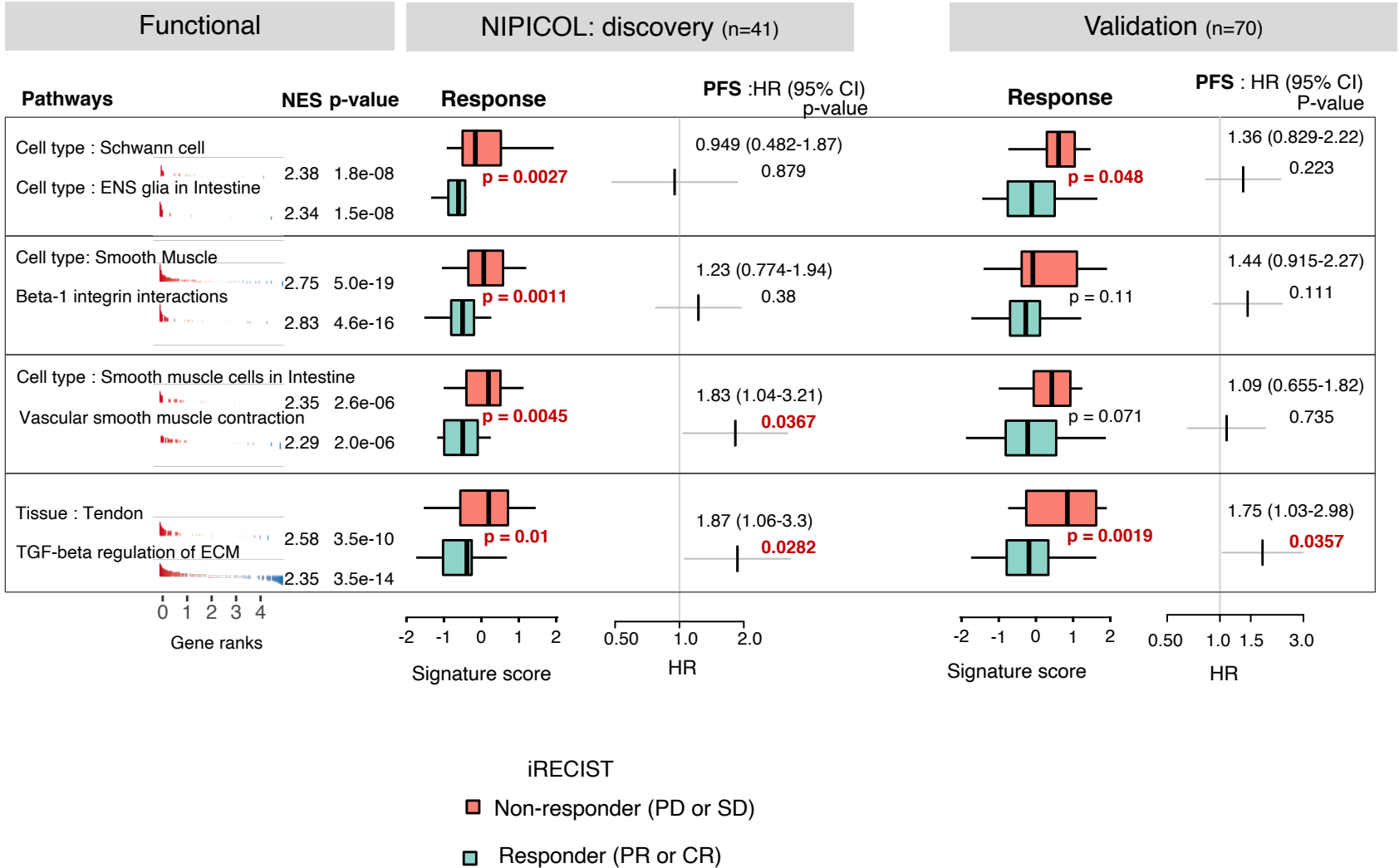
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

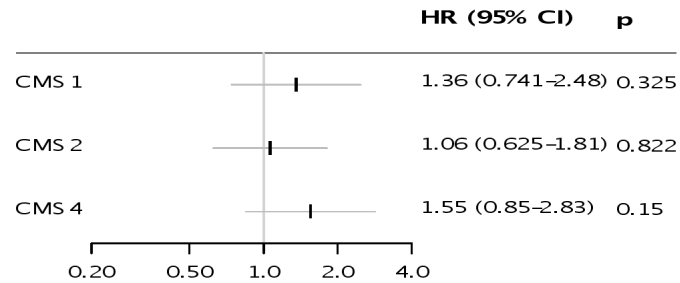
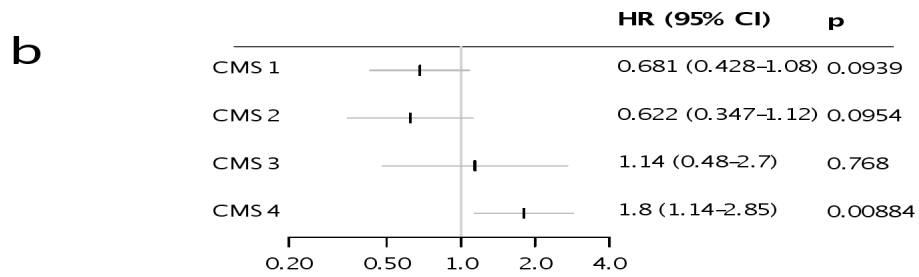
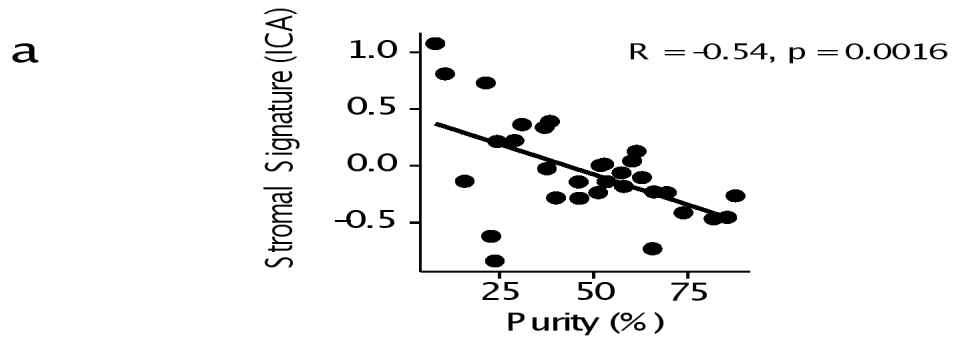
NIPICOL (discovery)

ImmunoMSI (validation)



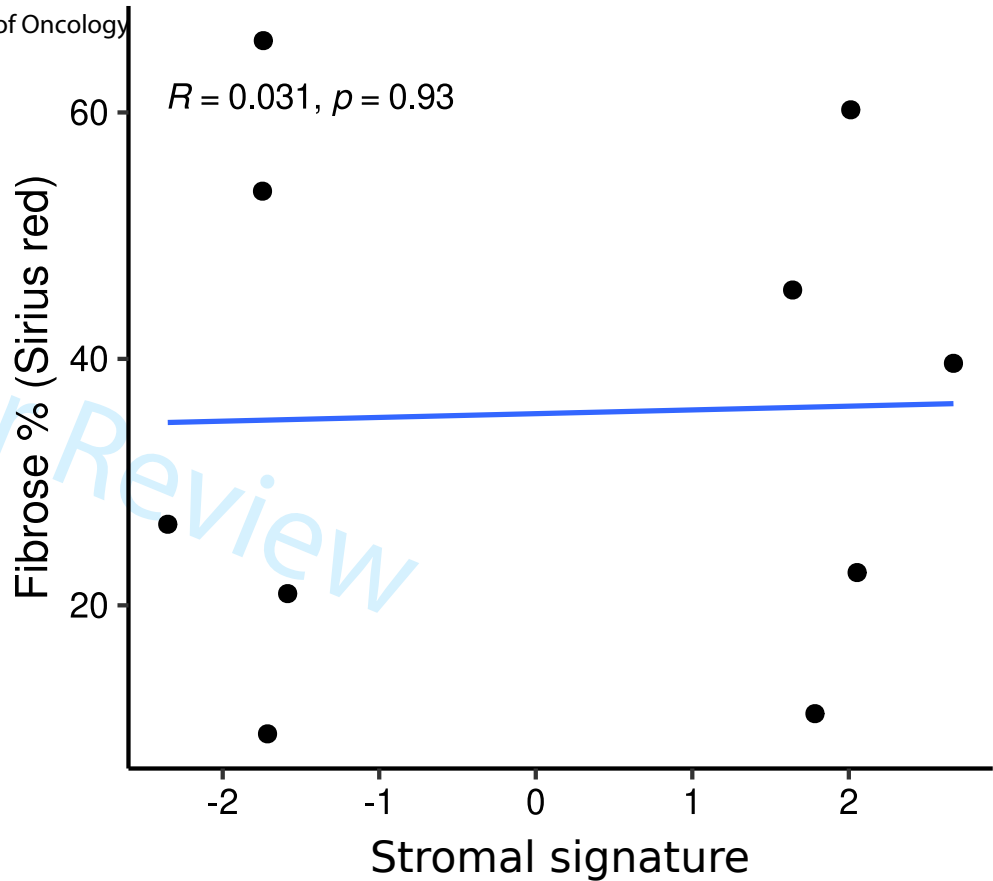
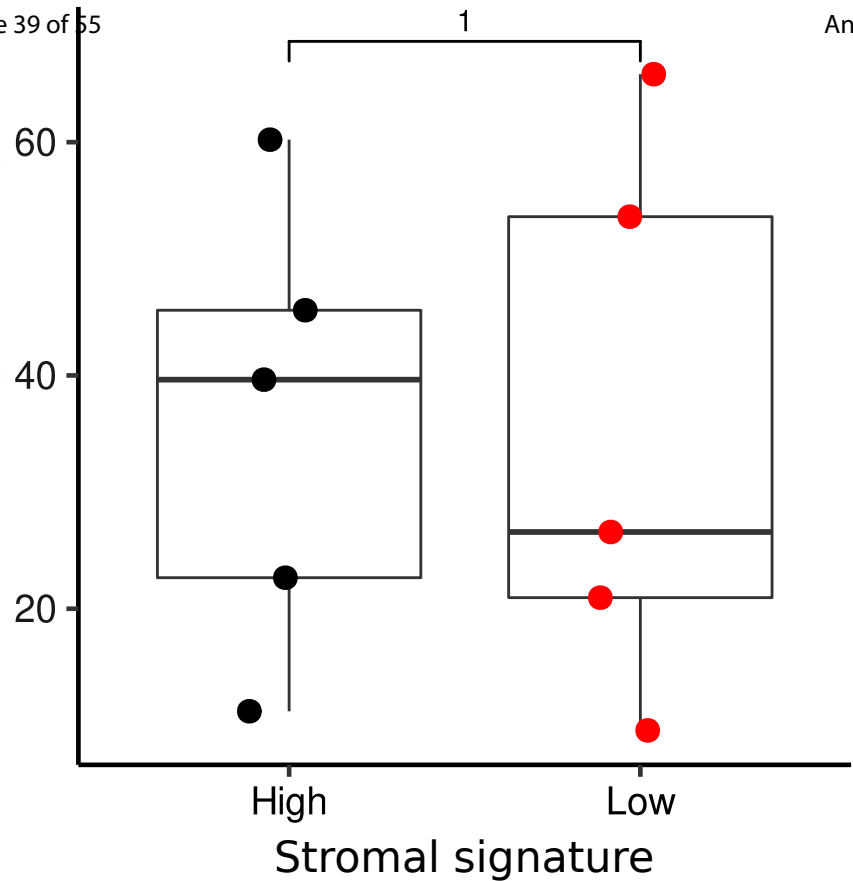
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

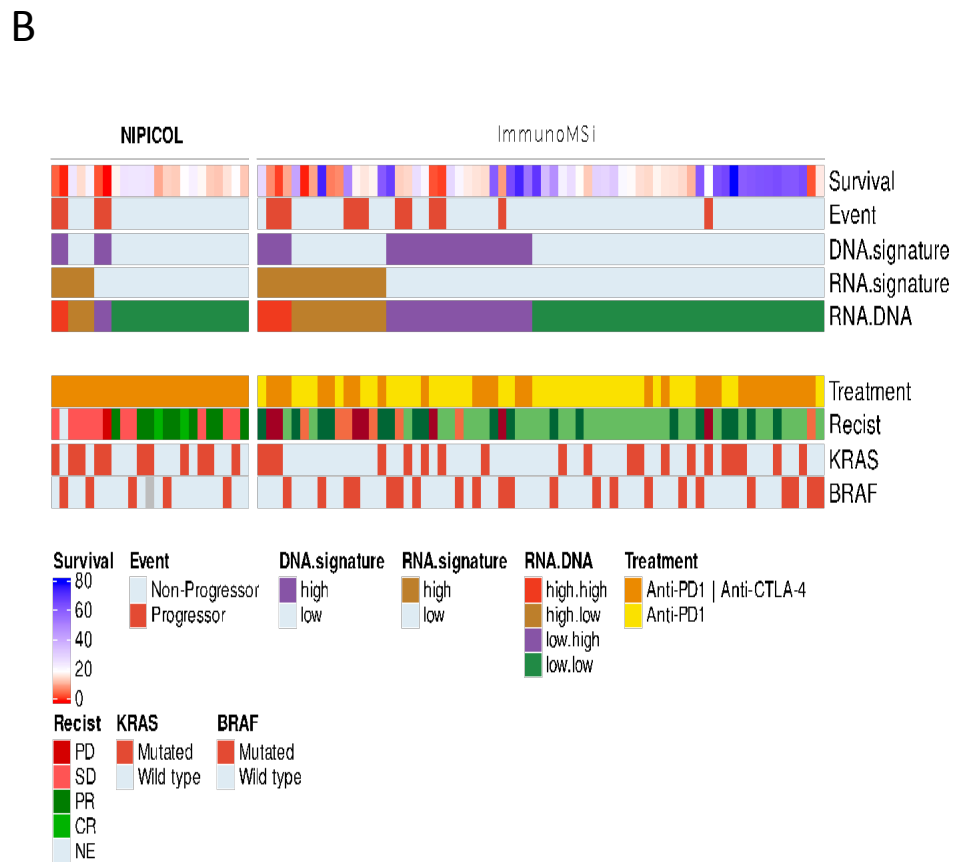
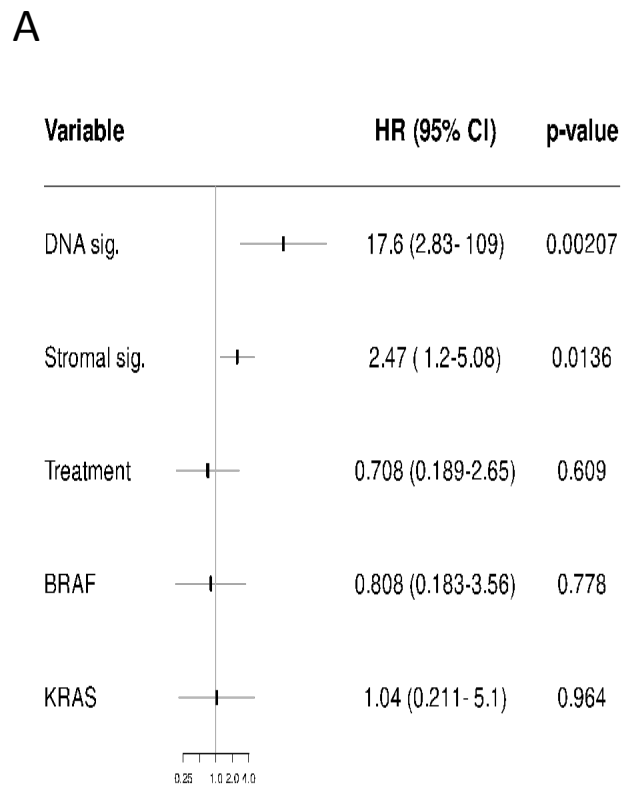


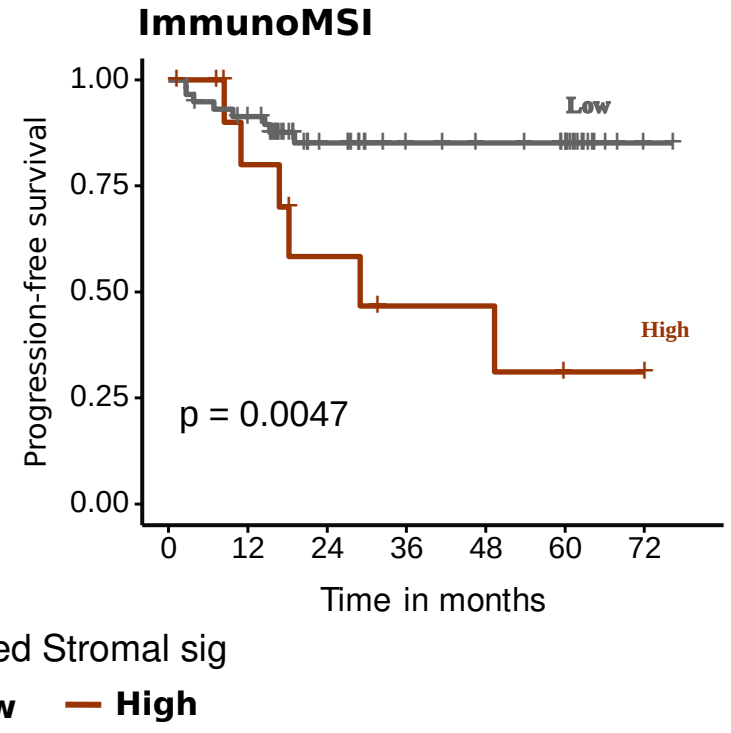
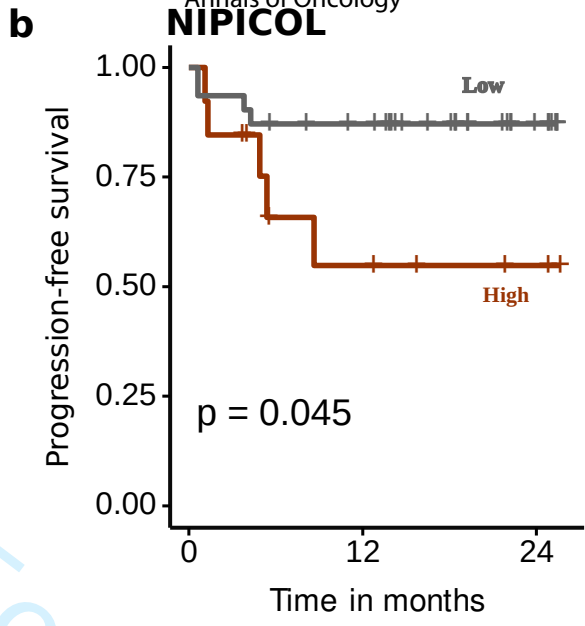
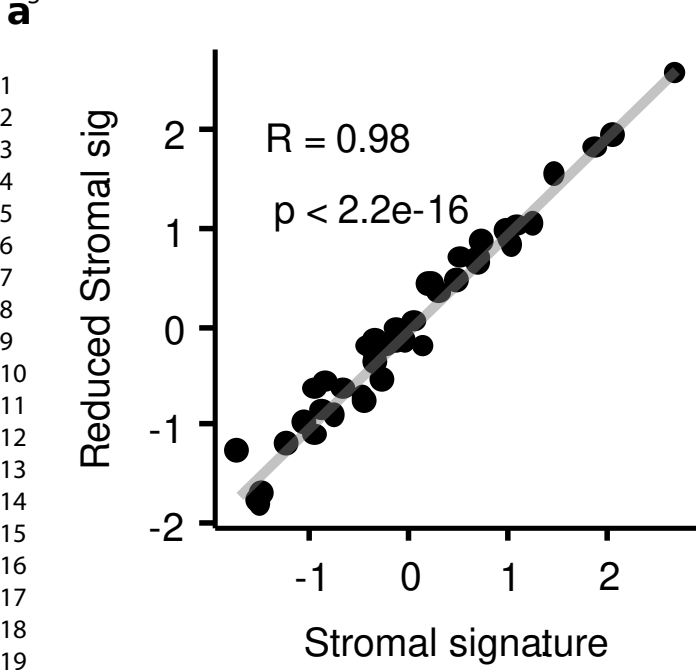


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Peer Review



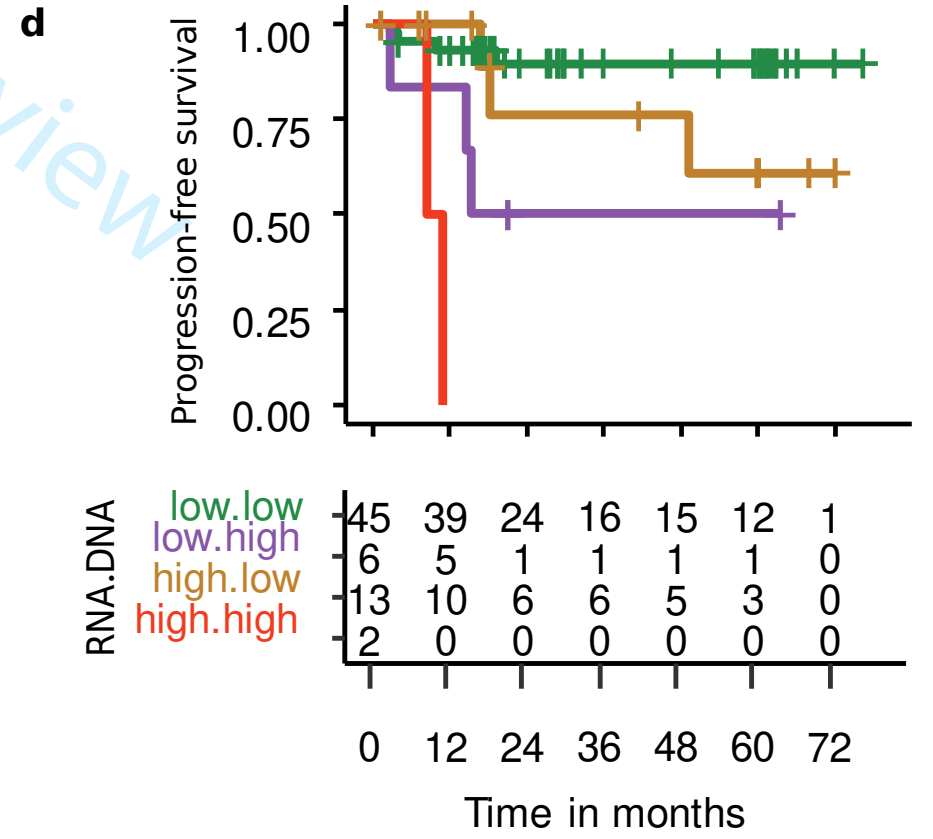




c

Continuous signatures in multivariate	HR (95% CI)	p-value
DNA sig.	13.9 (2.76-69.6)	0.00141
Reduced Stromal sig	65.7 (2.58-1.67e+03)	0.0113
Treatment	0.697 (0.206-2.35)	0.561

Multivariate cutoff signatures	HR (95% CI)	p-value
DNA sig. high	8.35 (2.49-27.9)	0.000579
Reduced Stromal sig. high	3.66 (1.01-13.3)	0.0489
Treatment	0.902 (0.261-3.12)	0.87



Characteristic	ImmunoMSI, N = 73¹	Nipicol, N = 47¹	p-value²
Age	59 (15)	54 (13)	0.058
Unknown	1	0	
Sex			0.87
Female	28 / 72 (39%)	19 / 47 (40%)	
Male	44 / 72 (61%)	28 / 47 (60%)	
Unknown	1	0	
ECOG performance score			0.24
≥1	36 / 71 (51%)	29 / 47 (62%)	
0	35 / 71 (49%)	18 / 47 (38%)	
Unknown	2	0	
Primary tumor location			0.49
Left colon	23 / 70 (33%)	18 / 46 (39%)	
Right colon	47 / 70 (67%)	28 / 46 (61%)	
Unknown	3	1	
BRAF			0.056
Mutated	23 / 72 (32%)	7 / 44 (16%)	
Wild type	49 / 72 (68%)	37 / 44 (84%)	
Unknown	1	3	
KRAS			0.029
Mutated	21 / 72 (29%)	23 / 47 (49%)	
Wild type	51 / 72 (71%)	24 / 47 (51%)	
Unknown	1	0	

Characteristic	ImmunoMSI, N = 73¹	Nipicol, N = 47¹	p-value²
NRAS			>0.99
Mutated	1 / 72 (1.4%)	1 / 43 (2.3%)	
Wild type	71 / 72 (99%)	42 / 43 (98%)	
Unknown	1	4	
Origin of MMR deficiency			0.82
Known germline mutation	20 / 30 (67%)	27 / 39 (69%)	
Sporadic	10 / 30 (33%)	12 / 39 (31%)	
Unknown	43	8	
Surgery of the primary tumor	71 / 72 (99%)	43 / 44 (98%)	>0.99
Unknown	1	3	
Number of metastatic sites			0.50
1	27 / 72 (38%)	14 / 47 (30%)	
2	26 / 72 (36%)	18 / 47 (38%)	
3	13 / 72 (18%)	10 / 47 (21%)	
4	6 / 72 (8.3%)	3 / 47 (6.4%)	
5	0 / 72 (0%)	2 / 47 (4.3%)	
Unknown	1	0	
Number of prior lines			0.006
0	9 / 72 (12%)	0 / 47 (0%)	
1	21 / 72 (29%)	8 / 47 (17%)	
2	32 / 72 (44%)	22 / 47 (47%)	
3	8 / 72 (11%)	12 / 47 (26%)	

Characteristic	ImmunoMSI, N = 73 ¹	Nipicol, N = 47 ¹	p-value ²
4	1 / 72 (1.4%)	3 / 47 (6.4%)	
5	1 / 72 (1.4%)	2 / 47 (4.3%)	
Unknown	1	0	
Type of immunotherapy			<0.001
Anti-PD1	43 / 72 (60%)	0 / 47 (0%)	
Anti-PD1 Anti-CTLA-4	29 / 72 (40%)	47 / 47 (100%)	
Unknown	1	0	
iRECIST			<0.001
CR	20 / 70 (29%)	2 / 44 (4.5%)	
PD	8 / 70 (11%)	3 / 44 (6.8%)	
PR	35 / 70 (50%)	15 / 44 (34%)	
SD	7 / 70 (10%)	24 / 44 (55%)	
Unknown	3	3	
Status			0.058
MSI	73 / 73 (100%)	44 / 47 (94%)	
MSS	0 / 73 (0%)	3 / 47 (6.4%)	
¹ Mean (SD); n / N (%)			
² Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test			

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ID	Cohort	Age	Sex	Status	DNA	RNA	Drug	Response	PFS	Event
IMSI_0150	Nipicol	50	Female	MSI	Exome	RNAseq	Anti-PD1 SD		25.29774	0
IMSI_0153	Nipicol	54	Male	MSI	Exome	RNAseq	Anti-PD1 PR		25.62628	0
IMSI_0188	Nipicol	49	Male	MSI	NA	RNAseq	Anti-PD1 PR		25.3963	0
IMSI_0195	Nipicol	39	Male	MSI	NA	RNAseq	Anti-PD1 PR		24.83778	0
IMSI_0181	Nipicol	64	Female	MSI	NA	RNAseq	Anti-PD1 SD		5.519507	0
IMSI_0187	Nipicol	38	Male	MSI	NA	RNAseq	Anti-PD1 SD		25.42916	0
IMSI_0180	Nipicol	58	Female	MSI	NA	RNAseq	Anti-PD1 SD		5.552361	0
IMSI_0151	Nipicol	32	Male	MSI	Exome	RNAseq	Anti-PD1 SD		25.03491	0
IMSI_0152	Nipicol	39	Male	MSI	Exome	RNAseq	Anti-PD1 PR		24.80493	0
IMSI_0179	Nipicol	60	Female	MSI	NA	RNAseq	Anti-PD1 SD		23.85216	0
IMSI_0146	Nipicol	63	Female	MSI	Exome	RNAseq	Anti-PD1 SD		24.80493	0
IMSI_0159	Nipicol	66	Male	MSI	NA	RNAseq	Anti-PD1 SD		8.640657	1
IMSI_0155	Nipicol	52	Male	MSI	Exome	RNAseq	Anti-PD1 SD		21.8152	0
IMSI_0189	Nipicol	63	Female	MSI	NA	RNAseq	Anti-PD1 SD		21.97947	0
IMSI_0154	Nipicol	49	Male	MSI	Exome	RNAseq	Anti-PD1 CR		10.97331	0
IMSI_0163	Nipicol	57	Female	MSI	Exome	RNAseq	Anti-PD1 SD		4.895277	1
IMSI_0156	Nipicol	63	Male	MSI	Exome	RNAseq	Anti-PD1 PR		14.68583	0
IMSI_0173	Nipicol	68	Male	MSI	NA	RNAseq	Anti-PD1 SD		21.61807	0
IMSI_0160	Nipicol	65	Male	MSI	Exome	RNAseq	Anti-PD1 SD		3.811088	1
IMSI_0165	Nipicol	54	Female	MSI	Exome	RNAseq	Anti-PD1 PR		22.20945	0
IMSI_0190	Nipicol	46	Male	MSI	NA	RNAseq	Anti-PD1 SD		22.2423	0
IMSI_0191	Nipicol	65	Male	MSI	NA	RNAseq	Anti-PD1 PR		13.96304	0
IMSI_0158	Nipicol	58	Female	MSI	Exome	RNAseq	Anti-PD1 PR		14.22587	0
IMSI_0194	Nipicol	72	Male	MSI	NA	RNAseq	Anti-PD1 PR		4.271047	1
IMSI_0175	Nipicol	44	Male	MSI	Exome	RNAseq	Anti-PD1 SD		19.25257	0
IMSI_0147	Nipicol	22	Male	MSI	Exome	RNAseq	Anti-PD1 SD		15.70431	0
IMSI_0162	Nipicol	66	Male	MSI	Exome	RNAseq	Anti-PD1 CR		19.21971	0
IMSI_0185	Nipicol	32	Male	MSI	Exome	RNAseq	Anti-PD1 #N/A		1.314168	1
IMSI_0161	Nipicol	50	Male	MSI	NA	RNAseq	Anti-PD1 SD		8.082136	0
IMSI_0148	Nipicol	31	Male	MSI	Exome	RNAseq	Anti-PD1 PR		18.06982	0
IMSI_0186	Nipicol	36	Male	MSI	Exome	RNAseq	Anti-PD1 PR		13.83162	0
IMSI_0184	Nipicol	36	Female	MSI	NA	RNAseq	Anti-PD1 #N/A		0.62423	1
IMSI_0169	Nipicol	33	Male	MSI	NA	RNAseq	Anti-PD1 PR		12.81314	0
IMSI_0167	Nipicol	61	Male	MSI	Exome	RNAseq	Anti-PD1 SD		18.3655	0
IMSI_0177	Nipicol	59	Female	MSI	NA	RNAseq	Anti-PD1 #N/A		1.117043	1
IMSI_0193	Nipicol	59	Male	MSI	NA	RNAseq	Anti-PD1 SD		18.43121	0
IMSI_0171	Nipicol	69	Male	MSI	NA	RNAseq	Anti-PD1 SD		3.679671	0
IMSI_0164	Nipicol	62	Female	MSI	Exome	RNAseq	Anti-PD1 PD		0.62423	1
IMSI_0178	Nipicol	76	Female	MSI	NA	RNAseq	Anti-PD1 SD		12.74743	0
IMSI_0166	Nipicol	61	Female	MSI	NA	RNAseq	Anti-PD1 SD		3.975359	0
IMSI_0192	Nipicol	37	Male	MSI	NA	RNAseq	Anti-PD1 SD		5.38809	1
IMSI_0172	Nipicol	56	Male	MSI	Exome	RNAseq	Anti-PD1 SD		16.45996	0
IMSI_0170	Nipicol	64	Female	MSI	Exome	RNAseq	Anti-PD1 PR		13.60164	0
IMSI_0168	Nipicol	77	Female	MSI	Exome	RNAseq	Anti-PD1 PR		14.22587	0
IMSI_0149	Nipicol	65	Female	MSS	Exome	RNAseq	Anti-PD1 PD		2.529774	1
IMSI_0157	Nipicol	51	Female	MSS	Exome	RNAseq	Anti-PD1 PD		4.13963	1
IMSI_0183	Nipicol	52	Female	MSS	Exome	RNAseq	Anti-PD1 PR		13.53593	0
RICKI-0701	ImmunoM	82	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		59.26667	0
RICKI-0531	ImmunoM	43	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		59.36667	0
RICKI-0601	ImmunoM	52	Male	MSI	Targeted S	RNAseq	Anti-PD1 PD		2.666667	1
RICKI-0011	ImmunoM	73	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		66.03333	0
RICKI-0541	ImmunoM	79	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		61.87097	0
RICKI-0091	ImmunoM	89	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		22.77419	0
RICKI-0711	ImmunoM	60	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		60.26667	0
RICKI-0081	ImmunoM	34	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		46.43333	0
RICKI-0641	ImmunoM	74	Male	MSI	Targeted S	RNAseq	Anti-PD1 PD		16.76667	1
RICKI-0721	ImmunoM	66	Male	MSI	Targeted S	RNAseq	Anti-PD1 SD		18.17857	0
RICKI-0551	ImmunoM	69	Female	MSI	Targeted S	RNAseq	Anti-PD1 SD		49.32258	1
RICKI-0021	ImmunoM	64	Female	MSI	Targeted S	RNAseq	Anti-PD1 PD		3.29	1
RICKI-0731	ImmunoM	48	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		61.13333	0
RICKI-0651	ImmunoM	64	Female	MSI	Targeted S	RNAseq	Anti-PD1 CR		64.36667	0
RICKI-0661	ImmunoM	65	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		76.26667	0
RICKI-0741	ImmunoM	64	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		71.8	0
RICKI-0671	ImmunoM	34	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		63.43333	0
RICKI-0751	ImmunoM	47	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		61.46667	0
RICKI-0481	ImmunoM	61	Male	MSI	Targeted S	RNAseq	Anti-PD1 PD		19.06452	1
RICKI-0691	ImmunoM	55	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		64.06667	0
RICKI-0761	ImmunoM	35	Female	MSI	Targeted S	RNAseq	Anti-PD1 CR		62.74194	0
RICKI-0621	ImmunoM	50	Male	MSI	Targeted S	RNAseq	Anti-PD1 PD		9.741935	1
RICKI-0771	ImmunoM	65	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		10.96667	1
RICKI-0491	ImmunoM	77	Female	MSI	Targeted S	RNAseq	Anti-PD1 CR		71.96774	0
RICKI-0781	ImmunoM	26	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		60.70968	0
RICKI-0461	ImmunoM	60	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		60.03333	0
RICKI-0791	ImmunoM	49	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		59.74194	0
RICKI-0801	ImmunoM	53	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		60.06667	0
RICKI-0811	ImmunoM	59	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		53.77419	0
RICKI-0821	ImmunoM	58	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		62.5	0
RICKI-0051	ImmunoM	33	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		41.36667	0
RICKI-0061	ImmunoM	45	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		35.83333	0
RICKI-0261	ImmunoM	62	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		32.4	0
RICKI-0071	ImmunoM	74	Female	MSI	Targeted S	RNAseq	Anti-PD1 SD		14.54839	1
RICKI-0031	ImmunoM	70	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		67.86667	0
RICKI-0611	ImmunoM	77	Male	MSI	Targeted S	RNAseq	Anti-PD1 PD		18.22581	1
RICKI-0171	ImmunoM	55	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		31.6	0
RICKI-0161	ImmunoM	50	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		29.73333	0
RICKI-0251	ImmunoM	63	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		28.73333	0
RICKI-0181	ImmunoM	67	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		29.56667	0
RICKI-0221	ImmunoM	65	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		2.642857	1
RICKI-0231	ImmunoM	61	Female	MSI	Targeted S	RNAseq	Anti-PD1 CR		28.8	0
RICKI-0241	ImmunoM	72	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		27.2	0
RICKI-0111	ImmunoM	51	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		27.53333	0
RICKI-0191	ImmunoM	48	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		27.1	0
RICKI-0211	ImmunoM	59	Female	MSI	Targeted S	RNAseq	Anti-PD1 PD		3.806452	1
RICKI-0281	ImmunoM	52	Male	MSI	Targeted S	RNAseq	Anti-PD1 SD		21.03333	0
RICKI-0201	ImmunoM	72	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		20.93333	0
RICKI-0291	ImmunoM	37	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		20.46667	0
RICKI-0121	ImmunoM	78	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		15.32258	1
RICKI-0131	ImmunoM	67	Female	MSI	Targeted S	RNAseq	Anti-PD1 CR		18.83333	0
RICKI-0841	ImmunoM	70	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		17.06667	0
RICKI-0311	ImmunoM	62	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		18.73333	0
RICKI-0371	ImmunoM	57	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		18.16667	0
RICKI-0381	ImmunoM	59	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		17.36667	0
RICKI-0391	ImmunoM	33	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		16.6	0
RICKI-0141	ImmunoM	61	Female	MSI	Targeted S	RNAseq	Anti-PD1 SD		1.193548	0
RICKI-0401	ImmunoM	28	Male	MSI	Targeted S	RNAseq	Anti-PD1 CR		16.36667	0
RICKI-0411	ImmunoM	61	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		16.33333	0
RICKI-0331	ImmunoM	78	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		16.03226	0
RICKI-0421	ImmunoM	65	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		15.76667	0
RICKI-0431	ImmunoM	75	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		15.56667	0
RICKI-0351	ImmunoM	58	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		15.36667	0
RICKI-0451	ImmunoM	58	Female	MSI	Targeted S	RNAseq	Anti-PD1 PD		8.433333	1
RICKI-0441	ImmunoM	43	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		11.93333	0
RICKI-0151	ImmunoM	83	Female	MSI	Targeted S	RNAseq	Anti-PD1 PR		13.96667	0
RICKI-0301	ImmunoM	36	Male	MSI	Targeted S	RNAseq	Anti-PD1 #N/A		6.866667	1
RICKI-0321	ImmunoM	40	Male	MSI	Targeted S	RNAseq	Anti-PD1 PR		10.4	0
RICKI-0511	ImmunoM	27	Male	MSI	Targeted S	RNAseq	Anti-PD1 SD		8.266667	0
RICKI-0501	ImmunoM	72	Female	MSI	Targeted S	RNAseq	Anti-PD1 CR		7.16129	0
RICKI-0831	ImmunoM	70	Male	MSI	Targeted S	RNAseq	Anti-PD1 SD		3.9	0
RICKI-0591	ImmunoM	78	Female	MSI	Targeted S	RNAseq	Anti-PD1 #N/A	#N/A	#N/A	#N/A

ENSid	Symbol	weight
ENSG000001	FNDC1	5.10420386
ENSG000001	SFRP2	4.30791092
ENSG000001	COMP	4.08019176
ENSG000001	AOC3	4.06685237
ENSG000001	EFEMP1	3.78175791
ENSG000001	MOGAT2	-3.7740728
ENSG000001	SFRP4	3.71450879
ENSG000001	MEDAG	3.68321604
ENSG000002	HNF1A-AS1	-3.6323767
ENSG000001	LRRCL5	3.58531426
ENSG000001	SYNC	3.55970288
ENSG000001	IHH	-3.4818844
ENSG000002	AC103702.2	-3.4495976
ENSG000002	MIR194-2HG	-3.4477355
ENSG000001	CHP2	-3.4388336
ENSG000001	STEAP4	3.42856296
ENSG000001	OPLAH	-3.4139494
ENSG000001	MYO1A	-3.3420832
ENSG000001	SSCSO	3.32262723
ENSG000001	DPT	3.28461284
ENSG000001	ESPN	-3.2834087
ENSG000001	DNMT1	-3.2508707
ENSG000001	MFAF5	3.20264084
ENSG000001	JCAD	3.17143233
ENSG000001	LHFPL6	3.1522329
ENSG000002	NEU4	-3.1504617
ENSG000002	AC132217.1	3.14969389
ENSG000001	CYP4F12	-3.1448367
ENSG000001	UCHL1	3.12997833
ENSG000001	GPR35	-3.1268892
ENSG000002	AL355987.4	-3.1196924
ENSG000001	GFPT2	3.1067951
ENSG000001	ASP	3.09936507
ENSG000000	YBX2	-3.0951517
ENSG000001	MOGAT3	-3.0917462
ENSG000001	FGF7	3.05626741
ENSG000002	LIX1	3.05311432
ENSG000002	AC104083.1	3.04291511
ENSG000001	MMP19	3.02492672
ENSG000000	IGF1	2.98795178
ENSG000001	COL10A1	2.98713455
ENSG000001	ITGBL1	2.98659956
ENSG000001	SLIT3	2.96293964
ENSG000001	ECM2	2.9601298
ENSG000000	DPEP1	-2.9377308
ENSG000001	TMTC1	2.93355859
ENSG000001	EMCN	2.90272468
ENSG000001	KIF12	-2.8889434
ENSG000000	SLCGA7	-2.8810662
ENSG000001	CYP2J2	-2.8798694
ENSG000002	GAS6-AS1	-2.8775443
ENSG000001	CYR61	2.85765177
ENSG000001	TSPAN2	2.85740009
ENSG000000	GUCY2C	-2.8539231
ENSG000001	CPED1	2.83495762
ENSG000001	RSPH1	-2.8306275
ENSG000002	TENM3	2.80218401
ENSG000001	MFAF4	2.80097456
ENSG000000	EML1	2.78458682
ENSG000001	LMOD1	2.78009958
ENSG000001	PDZRN3	2.77048747
ENSG000001	FABP1	-2.7664399
ENSG000001	LUM	2.76330963
ENSG000001	CES3	-2.742092
ENSG000001	DEGS2	-2.7379178
ENSG000001	CXCL12	2.73735607
ENSG000001	MGP	2.72643546
ENSG000001	TNFRSF11A	-2.7234544
ENSG000000	HSPB6	2.72059127
ENSG000001	ZNF296	-2.7164563
ENSG000001	PIGR	-2.7153683
ENSG000001	DDC	-2.7066467
ENSG000001	COLEC12	2.70112862
ENSG000001	ROR2	2.69878886
ENSG000000	CEACAM7	-2.6957512
ENSG000001	DUSP1	2.69013069
ENSG000001	KLRD1	2.68498717
ENSG000001	HOXB9	-2.6743332
ENSG000001	OLFM4	-2.6683462
ENSG000001	GGT6	-2.6586474
ENSG000001	PPP1R1B	-2.6554224
ENSG000001	CCNF	-2.6542216
ENSG000001	REEP6	-2.6406269
ENSG000001	RUNX2	2.63767694
ENSG000002	RNU6-1327P	-2.6359727
ENSG000001	MYO7B	-2.6349204
ENSG000000	LXN	2.63219126
ENSG000001	BDH1	-2.6320779
ENSG000000	PRKCZ	-2.630177
ENSG000000	PLA2G10	-2.6272585
ENSG000001	CRISPLD2	2.62520801
ENSG000001	SH3RF2	-2.6199242
ENSG000001	CPXM2	2.61602843
ENSG000001	A1CF	-2.6055727
ENSG000001	SULT1C2	-2.5911941
ENSG000002	SPIRE2	-2.5842921
ENSG000001	OSMR	2.57883696
ENSG000001	ANKS4B	-2.575791
ENSG000001	SLCGA20	-2.5589211
ENSG000001	TMEM102	-2.5543287
ENSG000001	HSD11B2	-2.5508296
ENSG000000	CFTF	-2.5507324
ENSG000001	EVPL	-2.5478452
ENSG000002	RN7S1559P	-2.5380794
ENSG000000	CCDC80	2.5284954
ENSG000001	CPNE7	-2.5164189
ENSG000001	TGFB3	2.51263919
ENSG000001	A4GALT	2.51250406
ENSG000002	VGLL3	2.51228151
ENSG000001	RAB17	-2.5109126
ENSG000000	PRSS8	-2.5103548
ENSG000001	COX7A1	2.50918775
ENSG000000	COX22	2.48249964
ENSG000001	CNN1	2.47735125
ENSG000001	CSMD2	2.47498771
ENSG000001	RHPN1	-2.4675944
ENSG000001	CYP11B1	2.45603284
ENSG000002	AC009133.1	-2.4556284
ENSG000002	STON1	2.453239
ENSG000001	FAM193B	-2.4480303
ENSG000001	CDH5	2.44664277
ENSG000000	ENTPD2	-2.4464431
ENSG000002	TNFSF12	2.4331992
ENSG000001	RNF43	-2.4330871
ENSG000001	BNC2	2.43236701
ENSG000001	TMEM45B	-2.426239
ENSG000002	LINC00623	2.42512842
ENSG000001	BOC	2.42172423
ENSG000001	TWIST1	2.42116864

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ENSG000002: CKMT1B	-2.4209972
ENSG000001: CACNA2D1	2.41522808
ENSG000000: PALM	2.40028821
ENSG000000: FAP	2.39438999
ENSG000000: COL16A1	2.39200067
ENSG000001: ADH1B	2.38962734
ENSG000000: MLXIPL	-2.3874269
ENSG000001: TNFAIP6	2.36927496
ENSG000001: FBLS5	2.36748176
ENSG000001: PRELP	2.36044837
ENSG000001: NOXO1	-2.3592309
ENSG000001: GIPC2	-2.3565795
ENSG000001: AZGP1	-2.3552531
ENSG000001: SAA1	2.35490426
ENSG000001: ANGPTL2	2.35244276
ENSG000000: STOML1	2.34418928
ENSG000001: SLC29A2	-2.341757
ENSG000002: BTBD19	2.33572824
ENSG000002: AC087388.1	-2.3264506
ENSG000001: LOX	2.32580601
ENSG000001: ZDHH23	-2.3129555
ENSG000001: PLACS	2.31118575
ENSG000001: ATP6V1C2	-2.3084971
ENSG000001: CSRN1P	2.30573514
ENSG000002: C14orf132	2.30110158
ENSG000001: LRP4	-2.2972655
ENSG000001: GUCY1A3	2.29672717
ENSG000001: THBS2	2.29125828
ENSG000002: DHR511	-2.2839635
ENSG000001: INCENP	-2.2834472
ENSG000001: ASCL2	-2.2830566
ENSG000001: RNF144B	2.27886073
ENSG000000: SEMA4G	-2.2785769
ENSG000001: SGPP2	-2.271301
ENSG000002: CPT1B	-2.2649018
ENSG000001: SPC25	-2.250969
ENSG000002: PHGR1	-2.2496352
ENSG000001: C10orf162	2.2469563
ENSG000002: MYO15B	-2.2407008
ENSG000001: ALDOB	-2.2399809
ENSG000001: MYB	-2.2368562
ENSG000001: PBLD	-2.2335698
ENSG000000: CDHR5	-2.2334101
ENSG000002: DNM3OS	2.23185767
ENSG000000: FKBP7	2.23117643
ENSG000001: ANO9	-2.2302804
ENSG000000: PARP3	-2.2300918
ENSG000000: SULT2B1	-2.2262162
ENSG000001: AGMAT	-2.2254478
ENSG000001: HMGCS2	-2.2204418
ENSG000001: MEI51	2.21438371
ENSG000001: CTGF	2.21025359
ENSG000001: PORDN1	2.21002783
ENSG000001: PDE10A	2.20695482
ENSG000002: LINC02086	-2.1984313
ENSG000000: IVD	-2.1955679
ENSG000001: PIEZO2	2.18800181
ENSG000001: TROAP	-2.1856571
ENSG000000: FBLN1	2.17547884
ENSG000001: C2orf82	-2.1711947
ENSG000001: RGMA	2.16953268
ENSG000001: OLFML2B	2.16885593
ENSG000001: GPC6	2.1653184
ENSG000002: AC018638.8	-2.1637824
ENSG000001: GLI3	2.16142102
ENSG000001: AMIGO2	2.15670404
ENSG000001: GPR78	2.15374608
ENSG000001: SSFN	2.15338107
ENSG000001: CTSK	2.15312225
ENSG000000: KCNQ1	-2.1526444
ENSG000000: ADAMTS2	2.15083602
ENSG000001: GDDP3	-2.150816
ENSG000001: CCSER1	-2.1499834
ENSG000001: PDSS1	-2.1449331
ENSG000001: CCL2	2.14424312
ENSG000000: FGR	2.14105989
ENSG000002: AC093840.1	-2.1362235
ENSG000001: DQX1	-2.1314835
ENSG000001: SELPLG	2.13018363
ENSG000001: VIPR1	-2.1300905
ENSG000001: ANPEP	-2.1287041
ENSG000002: AC018629.1	-2.1201524
ENSG000001: LCN2	-2.1194997
ENSG000001: C10orf10	2.11917274
ENSG000001: HOXB8	-2.1185241
ENSG000000: DCN	2.11845579
ENSG000002: AL359220.1	2.11668192
ENSG000001: CPZ	2.11349382
ENSG000001: TMPPRS2	-2.105955
ENSG000001: CTHRC1	2.0985844
ENSG000001: ADAM12	2.09492245
ENSG000001: MDFIC	2.0921297
ENSG000001: PCDH18	2.08884384
ENSG000000: ABCC9	2.08883494
ENSG000001: HOXB5	-2.0862379
ENSG000000: CHD1	-2.082894
ENSG000000: GRHL2	-2.0825277
ENSG000002: TRIM31	-2.0753583
ENSG000001: OLFML3	2.07502339
ENSG000000: BCS1L	-2.0699924
ENSG000001: DUSP7	2.06829642
ENSG000001: MNX1	-2.0681692
ENSG000002: CFI	2.06794984
ENSG000002: AC010503.4	-2.066242
ENSG000002: Metazoa_SRF	-2.0654099
ENSG000002: LINC01133	-2.0638474
ENSG000001: TJP3	-2.0638381
ENSG000001: LRG1	-2.0561147
ENSG000001: TMCS	-2.0496207
ENSG000001: GRASP	2.04825901
ENSG000001: CAVIN2	2.04445143
ENSG000001: ADGRL2	2.04185607
ENSG000001: IIGP	2.04035174
ENSG000001: TM4SF5	-2.032937
ENSG000001: GPR4	2.02614618
ENSG000002: C2orf72	-2.0235642
ENSG000001: HNF1A	-2.0217032
ENSG000001: CYP3A5	-2.0196008
ENSG000000: ATP2C2	-2.0193725
ENSG000001: AFAP1L1	2.01902945
ENSG000001: MST1	-2.0169282
ENSG000001: DTX3	2.01671556
ENSG000001: ADAMTS4	2.01491999
ENSG000001: SHH	-2.0132371
ENSG000001: BRIP1	-2.0122988
ENSG000001: KCNE4	2.01065693
ENSG000000: PLEKHG6	-2.0098157
ENSG000000: NGFR	2.00625839
ENSG000001: GRTP1	-2.0033983
ENSG000000: USH1C	-2.0017494

	*Target Identifier	Name
1	NM_005248.3	FGR
2	NM_000492.4	CFTR
3	NM_144617.3	HSPB6
4	NM_016429.4	COP22
5	NM_015982.4	YBX2
6	NM_153676.4	USH1C
7	NM_001291485.2	CEACAM7
8	NM_001384598.1	PLEKHG6
9	NM_203395.3	IYD
10	NM_032951.3	MLXIPL
11	NM_014228.5	SLC6A7
12	NM_001920.5	DCN
13	NM_001389466.1	DPEP1
14	NM_018397.5	CHDH
15	NM_000618.5	IGF1
16	NM_001003931.4	PARP3
17	NM_002773.5	PRSS8
18	NM_000218.3	KCNQ1
19	NM_203468.3	ENTPD2
20	NM_014861.4	ATP2C2
21	NM_002507.4	NGFR
22	NM_004809.5	STOML1
23	NM_002744.6	PRKCZ
24	NM_020297.4	ABCC9
25	NM_003561.3	PLA2G10
26	NM_004963.4	GUCY2C
27	NM_001079866.2	BCS1L
28	NM_006486.3	FBLN1
29	NM_004460.5	FAP
30	NM_024915.4	GRHL2
31	NM_001856.4	COL16A1
32	NM_014244.5	ADAMTS2
33	NM_177973.2	SULT2B1
34	NM_017893.4	SEMA4G
35	NM_021924.5	CDHR5
36	NM_002579.3	PALM
37	NM_006418.5	OLFM4
38	NM_024307.3	GDPD3
39	NM_031476.4	CRISPLD2
40	NM_001261841.2	TMC5
41	NM_001267560.2	TJP3
42	NM_000095.3	COMP
43	NM_000777.5	CYP3A5
44	NM_178176.4	MOGAT3
45	NM_003014.4	SFRP4
46	NM_000168.6	GLI3
47	NM_017680.5	ASPEN
48	NM_001393.4	ECM2
49	NM_199168.4	CXCL12
50	NM_017763.6	RNF43
51	NM_002982.4	CCL2
52	NM_003006.4	SELPLG
53	NM_000900.5	MGP
54	NM_001385079.1	PDE10A
55	NM_002317.7	LOX
56	NM_004624.4	VIPR1
57	NM_138393.4	REEP6
58	NM_001039348.3	EFEMP1
59	NM_024758.5	AGMAT
60	NM_001366006.2	ADGRL2
61	NM_001130173.2	MYB
62	NM_001901.4	CCN2
63	NM_003239.5	TGFB3
64	NM_024016.4	HOXB8
65	NM_002147.4	HOXB5
66	NM_004417.4	DUSP1
67	NM_000474.4	TWIST1
68	NM_005086.5	SSPN
69	NM_002429.6	MMP19
70	NM_000493.4	COL10A1
71	NM_007115.4	TNFAIP6
72	NM_001024630.4	RUNX2
73	NM_022449.4	RAB17
74	NM_024636.4	STEAP4
75	NM_017436.7	A4GALT
76	NM_001299.6	CNN1
77	NM_005515.4	MXN1
78	NM_005110.4	GFPT2
79	NM_003734.4	AOC3
80	NM_032192.4	PPP1R1B
81	NM_001082971.2	DDC
82	NM_005518.4	HMGCS2
83	NM_002262.5	KLRD1
84	NM_002334.4	LRP4
85	NM_000775.4	CYP2J2
86	NM_000545.8	HNF1A
87	NM_001166345.3	MDFIC
88	NM_032043.3	BRIP1
89	NM_000035.4	ALDOB
90	NM_001388308.1	KIF12
91	NM_182757.4	RNF144B

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

92	NM_000104.4	CYP1B1
93	NM_001370299.1	AMIGO2
94	NM_002345.4	LUM
95	NM_006329.4	FBLN5
96	NM_002009.4	FGF7
97	NM_003839.4	TNFRSF11A
98	NM_001554.5	CCN1
99	NM_000396.4	CTSK
100	NM_001039362.2	ATP6V1C2
101	NM_002398.3	MEIS1
102	NM_001378074.1	BOC
103	NM_003013.3	SFRP2
104	NM_003999.3	OSMR
105	NM_005564.5	LCN2
106	NM_014317.5	PDSS1
107	NM_014576.4	A1CF
108	NM_001288973.2	ADAM12
109	NM_001040694.2	INCENP
110	NM_080671.4	KCNE4
111	NM_020675.4	SPC25
112	NM_000722.4	CACNA2D1
113	NM_004181.5	UCHL1
114	NM_001378183.1	PIEZO2
115	NM_152550.4	SH3RF2
116	NM_052924.3	RHPN1
117	NM_130386.3	COLEC12
118	NM_005099.6	ADAMTS4
119	NM_080860.4	RSPH1
120	NM_001185.4	AZGP1
121	NM_203314.3	BDH1
122	NM_001864.4	COX7A1
123	NM_181711.4	TAMALIN
124	NM_001761.3	CCNF
125	NM_002644.4	PIGR
126	NM_152386.4	SGPP2
127	NM_012134.3	LMOD1
128	NM_002181.4	IHH
129	NM_001443.3	FABP1
130	NM_020208.4	SLC6A20
131	NM_001947.4	DUSP7
132	NM_001130682.3	GUCY1A1
133	NM_000193.4	SHH
134	NM_138455.4	CTHRC1
135	NM_025098.4	MOGAT2
136	NM_002404.3	MFAP4
137	NM_001150.3	ANPEP
138	NM_005379.4	MYO1A
139	NM_022097.4	CHP2
140	NM_001288702.2	GGT6
141	NM_001988.4	EVPL
142	NM_206918.3	DEGS2
143	NM_004657.6	CAVIN2
144	NM_004560.4	ROR2
145	NM_001393586.1	MYO7B
146	NM_024017.5	HOXB9
147	NM_052972.3	LRG1
148	NM_024922.6	CES3
149	NM_199161.5	SAA1
150	NM_020998.4	MST1
151	NM_001532.3	SLC29A2
152	NM_000196.4	HSD11B2
153	NM_005282.3	GPR4
154	NM_178502.4	DTX3
155	NM_005301.5	GPR35
156	NM_153636.3	CPNE7
157	NM_017570.5	OPLAH
158	NM_001795.5	CDH5
159	NM_001144950.2	SSCS5
160	NM_020211.3	RGMA
161	NM_005708.5	GPC6
162	NM_005656.4	TMPRSS2
163	NM_003062.4	SLIT3
164	NM_001012302.3	ANO9
165	NM_023944.4	CYP4F12
166	NM_003247.5	THBS2
167	NM_031475.3	ESPN
168	NM_001377530.1	DMBT1
169	NM_002725.4	PRELP
170	NM_172167.3	NOXO1
171	NM_000668.6	ADH1B
172	NM_003480.4	MFAP5
173	NM_001056.4	SULT1C2
174	NM_004791.3	ITGBL1
175	NM_001167600.3	NEU4
176	NM_007028.5	TRIM31
177	NM_032451.2	SPIRE2
178	NM_000204.5	CFI
179	NM_152246.3	CPT1B
180	NM_001375484.1	CKMT1B
181	NM_003809.3	TNFSF12
182	NM_006873.4	STON1

SUPPLEMENTARY MATERIALS AND METHODS

Sample preparation and sequencing

RNA sequencing was performed on the same platform for all patient cohorts included in this work, whether or not they were treated with ICI. For the DNA analyses, all exome sequencing was performed with comparable qualitative parameters (e.g., ~50X depth for normal samples and ~200X depth for tumor samples). The exome profiles were produced by the Integragen platform. For the C2 cohort (targeted NGS panel, ICM platform), depth of coverage was likely similar, comprised between 50X and 300X. For the C1 cohort, genomic DNA isolated from formalin- fixed paraffin- embedded (FFPE) were captured using Twist Human Core Exome Enrichment System (Twist Bioscience) + IntegraGen Custom as previously described ¹. For the C2 cohort, hybridization probes and extension primers were designed using Roche's HyperDesign Tool for NGS Target Enrichment process. Sequence capture, enrichment and elution were performed according to manufacturer's instruction and protocols without modification except for library preparation performed with NEBNext® Ultra II kit (New England Biolabs®). DNA samples were then sequenced on an Illumina NovaSeq as Paired-end 100 reads. The generated reads were mapped to the reference genome hg38 (GRCh38).

Exome Analysis and Mutational load

Variant calling for the identification of SNVs (Single Nucleotide Variations) and small insertions/deletions (up to 20bp), was performed via the Broad Institute's GATK Haplotype Caller GVCF tool (GATK3.8.1) for constitutional DNA and via the Broad Institute's MuTect tool (2.0, --max_alt_alleles_in_normal_count=2; --max_alt_allele_in_normal_fraction=0.04) for somatic DNA. A two-sided Fisher's Exact Test is applied after MuTect2 variant calling to improve filtering of variants with strand bias. The tumor mutational burden (TMB) is calculated by dividing the number of somatic mutations by the number of bases having a depth greater than 10. The somatic mutations used for the mutational load were filtered as follows: Somatic score ≥ 3 ,

1
2
3 Mutated Allele Frequency in Tumor tissue $\geq 5\%$, Mutated Allele Count in Tumor tissue
4 ≥ 3 , Mutated Allele Frequency in Constitutional tissue $< 4\%$.
5
6
7
8
9

10 MHC class-I neoepitopes prediction

11
12 Neopeptides prediction was performed using *Ideation@SiRIC* pipeline which
13 combines several software. First, HLA typing for MHC class-I genes was carried out
14 using seq2HLA ² software for all 26 normal-wes fastq files, with default parameters.
15 Next, Somatic mutation filtered VCF files from Mutect2 were annotated by Variant
16 Effect Predictor (Version 99) with default parameter, except for the using of 'Frameshift'
17 and 'wild-type' plugins offer by pVACseq from pVACtools suite ³. After annotation, the
18 variant items that lead to peptide changes were extracted out for downstream analysis.
19 vcf-expression-annotator from VAtools v5.0.1 were then used to add transcripts
20 expression levels (using 3' RNA sequencing) for each variant. Annotated non-
21 synonymous mutations, frameshift variants and sequencing-based information as well
22 as HLA class-I gene typing results inferred by seq2HLA were fed into pVACseq for
23 neoantigen prediction. For each pVACseq run, epitope prediction was done by both
24 NetMHC ⁴ and NetMHCpan ⁵ algorithms packed in pVACseq toolkit, epitope length
25 was set to 8–10 with default parameters used for all other settings. At last, we applied
26 coverage and expression-based filters besides median binding affinity filtering criteria
27 on the predicted neoepitopes. We filtered neoepitope with tumor DNA VAF $\leq 5\%$,
28 along with TPM ≥ 1 for transcript level expression. Afterward, we kept only
29 neoepitopes with TSL (Transcript Support Level) equal to 1 and median affinity binding
30 ≤ 500 . To calculate the number of neoepitopes per tumor, we sum up unique filtered
31 mutants peptides sequences per tumor.
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 Feature selection and data pre-processing

50
51 To reduce our dataset (over 279 positions explored) to the most relevant positions,
52 we decided to keep only the positions with a Cox regression p -value in Nipicol lower
53 than 0.05 ⁶. We also decided to only focus on variations in microsatellite regions.
54 Finally, we have applied a filter to the maximum rate of missing data allowed in a
55 position in Nipicol. To do so, we firstly calculated this value in other available WES
56 cohorts of MSI CRC, i.e., the public cohort TCGA and two other large private cohorts
57
58
59
60

1
2
3 we have already sequenced and analyzed in the lab (data not shown). This allows us
4 to exclude positions that are generally difficult to sequence. Secondly, we did the same
5 on the Nipicol cohort to exclude the remaining mis-sequenced positions. We have
6 chosen to allow a maximum of 5% of missing data per MS over the external WES or
7 Nipicol. With those filters we selected 19 microsatellite variations of interest.
8
9
10
11

12 In order to train a MNB model on our DNA sequencing data, we had process missing
13 sequencing data at certain locations for some patients. Indeed, MNB does not support
14 missing data in its learning or prediction. As a result, the missing data must be imputed.
15 We imputed missing data using by using a discrete uniform distribution drawing wild-
16 type or mutated based on the mutation frequency of the microsatellite.
17
18
19
20
21
22
23

24 **DNA signatures : mutation count and Multinomial Naives Bayes**

25
26 To calculate normalized mutation count we firstly had to take into account the
27 impact of the mutation on the ICI response. To do so, we inverted the status of the
28 microsatellite mutation associated with a good response to ICI. As a results we
29 obtained a matrix containing for each patient the genotype effect for each microsatellite
30 (0 for beneficial and 1 for deleterious). Then, for each patient, we summed up the
31 microsatellite status and divided it by the number of microsatellite sequenced.
32
33
34
35
36

37 For the MNB, we used a matrix containing the mutation status for each microsatellite
38 in every Nipicol patient (1 for mutated, 0 for wild-type) as the input values and the iPFS
39 status as the binary target values. We obtained a model that was able to predict the
40 progression risk of a patient taking into account its sequencing on the 19 microsatellites
41 investigated.
42
43
44
45

46 To validate our models, we calculated the mutation count and predicted the
47 progression risk for each ImmunoMSI patients. Then we ran an univariate Cox
48 regression model to investigate the association between the survival and those new
49 signatures.
50
51
52
53
54
55

56 **Transcriptome profile generation and analysis**

57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Total RNA was extracted for every tumor sample using QIAGEN Allprep protocol (ref 10.2144/000113829). 150ng of total RNA was used to generate RNAseq sequencing library using Lexogen Quantseq 3' FWD kit and sequenced by the Institut du cerveau iGenseq platform on a NovaSeq 6000 aiming for a minimum of 10 million reads. Raw RNAseq reads were mapped to the human genome (Ensembl GRCH38) and Ensembl's reference transcriptome using STAR. Gene counts were obtained using FeatureCount, normalized by an UpperQuartile procedure and logged on a base 2.

Gene-set definitions of pathways were taken from the Molecular Signature Database (MSigDB v7.5.1) as well as the BioPlanet pathways database (2019) which was downloaded from the EnrichR website. Pathway-level activity was inferred from gene expression levels using the GSVA method with a Gaussian kernel. Based on the ImmuneDeconv evaluation and implementation, five methods for tumor microenvironment quantification were used: XCell, EPIC, Cibersort, MCPcounter and Estimate. Multiple hypotheses testing correction using Benjamini-Hochberg's to control False Discovery Rate (FDR) was only applied to select significant events using an alpha 5% to reject H0.

Independent Component Analysis (ICA) was performed on the sample-wise centered gene expression values of the 5,000 genes with the highest average and variable (as measured by standard deviation) expression. The JADE (joint approximation diagonalization of Eigen-matrices) algorithm was applied to extract 10 components as previously described ⁷. The ICA components were projected on other transcriptomic dataset by first selecting genes in common in the reference ICA (computed from the C1 cohort) and in the new transcriptomic dataset. The cross-product of the gene-weight inverted matrix was used as the projection of a new dataset, resulting in 10 scores (one per ICA component) for each new sample.

The intra-tumor proportion of the consensus molecular subtypes of colorectal cancer were estimated using the centroids of the original study ⁸ and the WISP deconvolution method (cit-bioinfo.github.io/WISP/).

- 1 Gnrirke A, Melnikov A, Maguire J et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol* 2009; 27 (2): 182-189.
- 2 Boegel S, Lower M, Schafer M et al. HLA typing from RNA-Seq sequence reads. *Genome Med* 2012; 4 (12): 102.
- 3 Hundal J, Kiwala S, McMichael J et al. pVACtools: A Computational Toolkit to Identify and Visualize Cancer Neoantigens. *Cancer Immunol Res* 2020; 8 (3): 409-420.
- 4 Lundegaard C, Lamberth K, Harndahl M et al. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. *Nucleic Acids Res* 2008; 36 (Web Server issue): W509-512.
- 5 Reynisson B, Alvarez B, Paul S et al. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* 2020; 48 (W1): W449-W454.
- 6 Polsterl S, Conjeti S, Navab N et al. Survival analysis for high-dimensional, heterogeneous medical data: Exploring feature extraction as an alternative to feature selection. *Artif Intell Med* 2016; 72: 1-11.
- 7 Puleo F, Nicolle R, Blum Y et al. Stratification of Pancreatic Ductal Adenocarcinomas Based on Tumor and Microenvironment Features. *Gastroenterology* 2018; 155 (6): 1999-2013 e1993.
- 8 Guinney J, Dienstmann R, Wang X et al. The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015; 21 (11): 1350-1356.

1
2
3
4 **Supplementary Table S1: Patients description from C1 and C2 cohorts with**
5 **corresponding clinical data**

6
7 ECOG, Eastern Cooperative Oncology Group performance-status scores range from
8 0 to 5, with higher scores indicating greater disability. CR, complete response; PR,
9 partial response; SD, stable disease; PD, progressive disease.
10
11
12

13
14 **Supplementary Table S2: NIPICOL and ImmunoMSI cohorts with clinical data**
15 **including survival data**

16
17
18
19 **Supplementary Table S3: List of the genes most contributing to the stromal signature**

20
21
22
23 **Supplementary Table S4: List of the genes (n=182 genes) in the reduced stromal**
24 **signature**

25
26
27
28 **Supplementary Figure S1. Correlations between MSISensor, MSICare and other**
29 **genomic instability indexes, analysis of their clinical relevance to predict ICI resistance**
30 **in MSIImCRC patients**

31
32
33
34
35 **Supplementary Figure S2. Progression free survival (iPFS) curve according to non-**
36 **repetitive Gene variant status**

37
38 Kaplan-Meier survival curves plotted for progression-free survival (iPFS) according to
39 some Genes mutation status (Wild type are indicated in grey and mutated in dark-red).
40
41
42

43
44 **Supplementary Figure S3: Predictive nature of this 19-plex identified DNA signature**

45
46 Analysis based on the concomitant analysis of their somatic variations in tumor DNA
47 as compared to matched normal DNA by simply calculating for each patient the mean
48 count of all the mutations in those MS.
49
50
51

52
53 **Supplementary Figure S4: Transcriptome analysis**

54
55
56 **A-B) Univariate analyses of the association between cell population quantification**
57 **estimation or single-gene expression levels and the objective response to ICI as**
58 **evaluated by iRECIST and iPFS in the two cohorts C1 (a) and C2 (b).**
59
60

1
2
3
4 **(C-H)** Distribution of Cox's proportional hazard p -value measuring the association of
5 tumor microenvironment signatures (c,f), Pathway's estimated activity (d,g) or single-
6 gene expression levels (e,h) with progression-free survival in C1 and in either the entire
7 C2 (c,d,e) or only in the subset of C2 in which patients received the same combination
8 of ICI than in C1 (f,g,h). In the plots, p -values are signed so that association with
9 increased risk ($HR > 1$) are positive and association with decreased risk ($HR < 1$) are
10 negative.
11
12
13
14
15
16

17 **Supplementary Figure S5. Testing association of pathway activity to objective**
18 **response as evaluated by iRECIST and progression free survival by iRECIST (iPFS)**

19 Univariate association of pathway activity estimated by GSVA in C1 NIPICOL cohort
20 (left) and C2 ImmunoMSI cohort (right). Gene set defining pathways were taken
21 Bioplanet collection by default, or Elsevier pathway collection and BIOCARTA when
22 specified.
23
24
25
26
27

28
29 **Supplementary Figure S6. Characterization of the four RNA components defined by**
30 **ICA in C1 NIPICOL**

31 Four RNA components defined by ICA were significantly associated with objective
32 response defined by iRECIST. From left to right: Pathway enrichment (NES:
33 normalized enrichment score), C1 iRECIST objective response, univariate association
34 with iPFS in C1, association with objective response defined by iRECIST in C2 and
35 univariate association with iPFS in C2. The boxes go from the upper to the lower
36 quartiles, and the whiskers go from the minimum to the maximum value, the line in the
37 box corresponds to the median (C1: $n=41$; C2: $n = 70$).
38
39
40
41
42
43
44
45

46
47 **Supplementary Figure S7. RNA signatures**

48 A) Correlation of the stromal signature, defined by ICA, with tumor cellularity estimated
49 by exome sequencing.
50

51 B) Univariate association of the intra-tumoral proportion of CMS with progression-free
52 survival by iRECIST.
53
54
55

56
57 **Supplementary Figure S8. Quantification of Sirius Red staining in MSI mCRC samples**
58 **with low or high stromal RNA signature**
59
60

1
2
3
4 A) Box plot of the percentage of fibrosis assessed by automated morphometry and
5 according to the stromal signature (high >0 , low <0). The boxes go from the upper to
6 the lower quartiles and the line in the box corresponds to the median.
7

8 B) Correlations between the percentage of fibrosis assessed by automated
9 morphometry and the stromal signature.
10
11
12
13

14 **Supplementary Figure S9: Multivariate analysis of both signatures in patients**

15
16

17 A) Forest plot of the multivariate Cox regression with progression-free survival in cohort
18 C2, representing hazard ratio (HR < 1 : beneficial, HR = 1 No effect, HR > 1 :
19 deleterious) with its corresponding p.value. Included features are DNA signature (low
20 to high risk), stromal RNA signature (low to high fibrose score), BRAF, KRAS (wild-
21 type vs mutated), and treatment type (mono vs. combo).
22
23

24 B) Heatmap representation, in C1 (n=23) and C2 (n=66), of the DNA and RNA models
25 predictions in regard with the patients iPFS, iRECIST status and the treatment (combo-
26 therapy or mono-therapy) (27).
27
28
29
30
31

32 **Supplementary Figure S10: Reduced RNA stromal signature**

33

34 A) Spearman correlation between the original transcriptome-wide component and the
35 reduced signature.
36

37 B) Kaplan-Meier curves of iPFS using reduced RNA signature (n=182 genes) and the
38 same cut-off as with the original signature in Nipicol (C1) and ImmunoMSI (C2).
39

40 C) Forestplot of the reduced RNA stromal signature and the DNA signature as well as
41 the treatment type considering the continuous value (upper panel) or by discretizing
42 the values (lower panel) of the signatures.
43
44

45 D) Kaplan-Meier estimates of iPFS as a function of the intensity combination of the
46 reduced RNA signature and the DNA signature (n=66). High RNA and DNA signatures
47 are shown in red. Low RNA and DNA signatures are shown in green. High RNA
48 signature and weak DNA signature are shown in orange and low RNA signature and
49 high DNA signature are shown in purple.
50
51
52
53
54
55
56
57
58
59
60