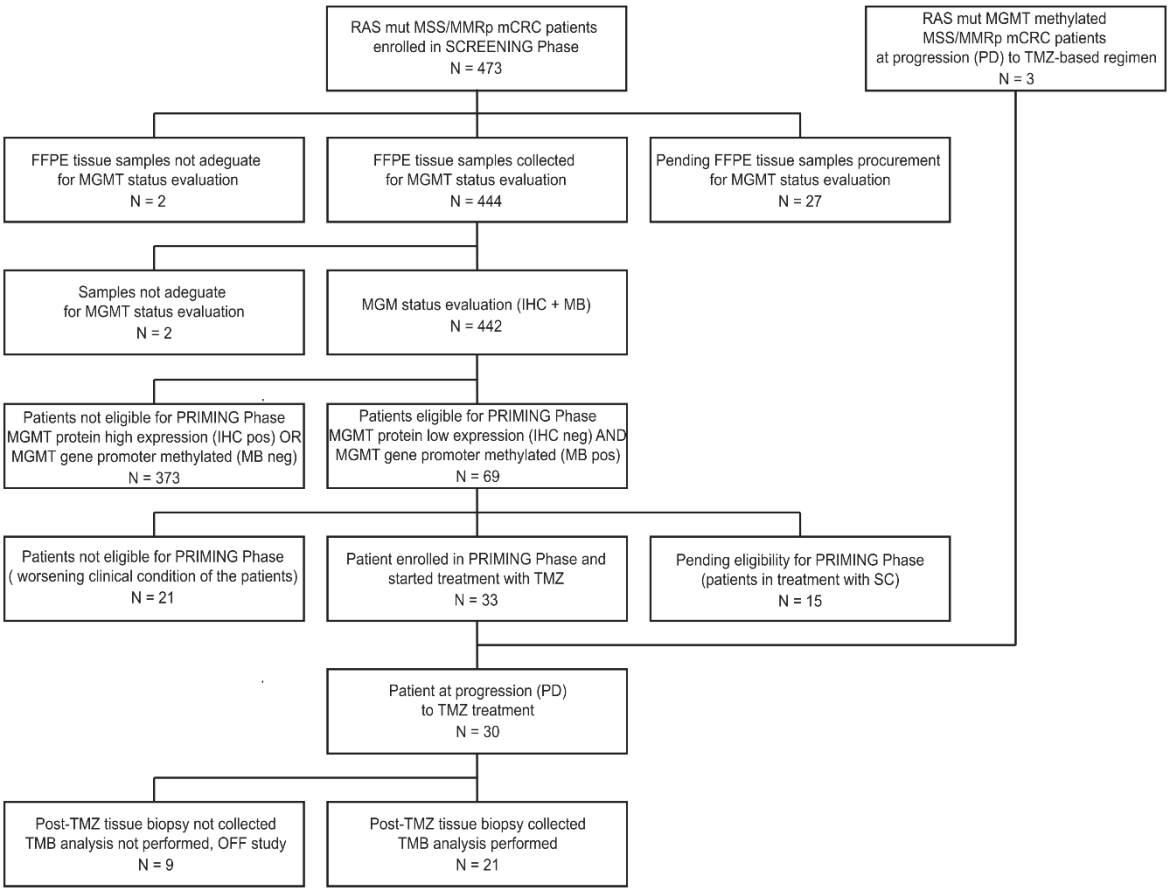
Supplementary Data Information and description

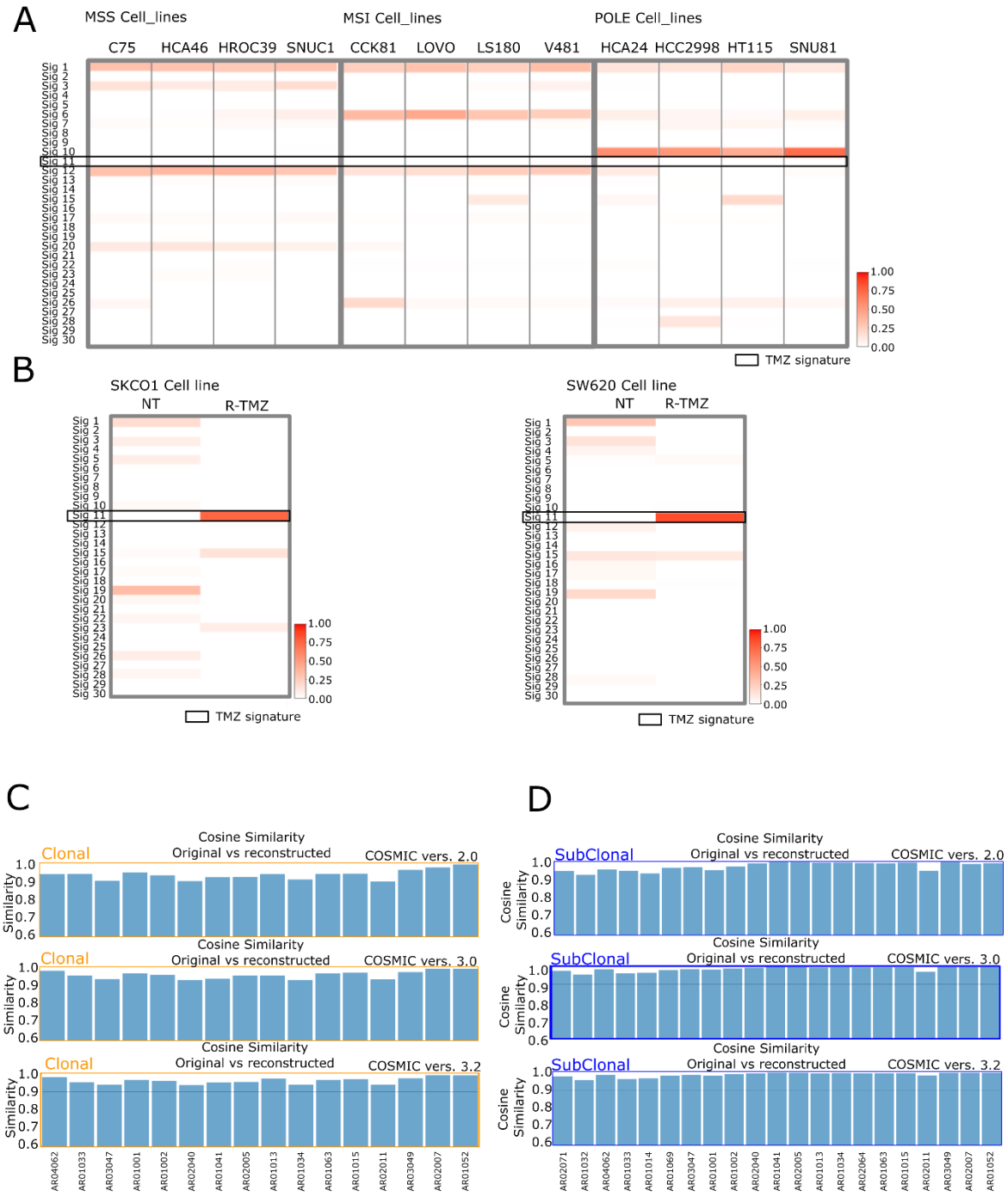
Supplementary Figure S1



Supplementary Figure S1:

CONSORT diagram of patients screened for the ARETHUSA trial (N=473), and eligible for the temozolomide PRIMING (N=69) and evaluated according to IMMUNOTHERAPY eligibility criteria (N = 30) as described in this study. *N = Numbers; TMZ = Temozolomide; PD=progressive disease; MSS = microsatellite stable; MMRp = mismatch repair proficient; mut = mutant; FFPE = Formalin-Fixed Paraffin-Embedded; mCRC = metastatic Colorectal cancer; IHC=immunohistochemistry; MB=methyl beaming; pos = positive; neg = negative; TMB = Tumor Mutational Burden; SC =Standard Care; O6-Methylguanine-DNA-methyltransferase (MGMT).*

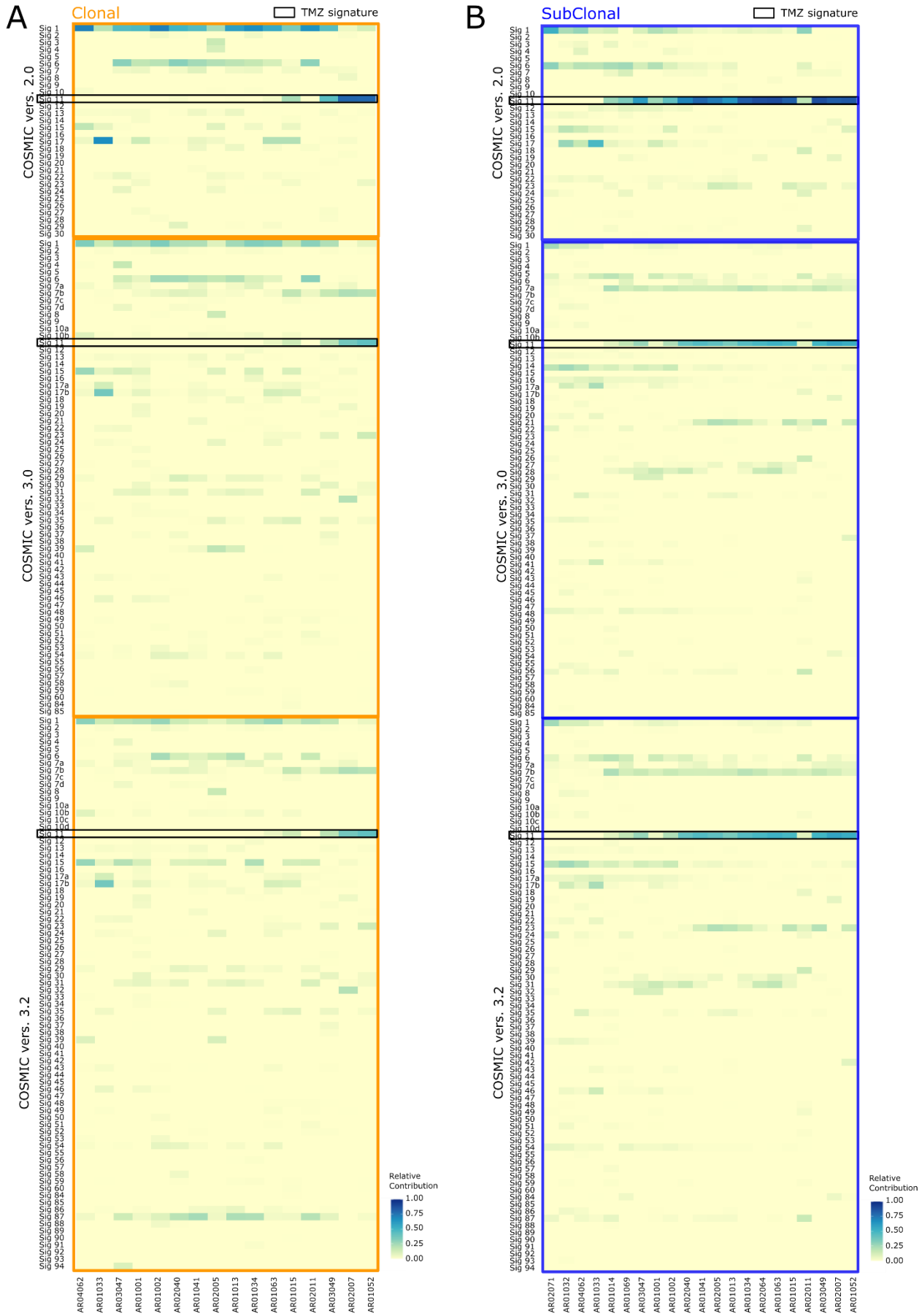
Supplementary Figure S2



Supplementary Figure S2:
A) Profile of mutational signatures in a panel of 12 CRC cell lines according to the indicated genetic features. **B)** Profile of mutational signatures in parental SKCO1 and SW620 cell lines and in their R-TMZ resistant counterparts**. C-D)** Comparison of cosine similarity calculated between original vs reconstructed mutational profiles of tissues post TMZ. Three different reference databases were used: COSMIC version 2.0 (top), COSMIC vers 3.0 (center) and COSMIC vers 3.2 (bottom) **C)** Analysis was performed using only clonal mutations (fractional abundance ≥ 10%). In five cases (AR02071, AR01032, AR01014, AR01069 and AR02064) the cosine similarity was lower than 0.9. **D)** Analysis was performed using both

clonal and subclonal mutations (fractional abundance ≥ 1%). *vers. = version; TMZ= Temozolomide; NT= not treated; R-TMZ= resistant to TMZ; Sig= Signature; MSS=microsatellite stable; MSI= microsatellite instable; DNA polymerase epsilon (POLE) mutations*.

Supplementary Figure S3

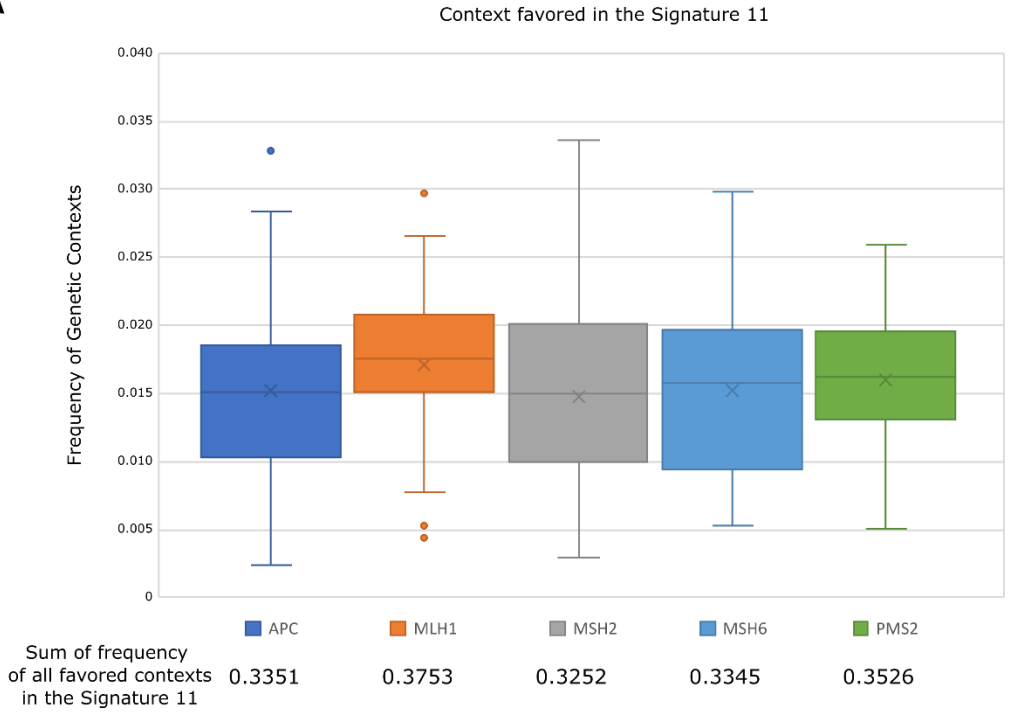


Supplementary Figure S3:

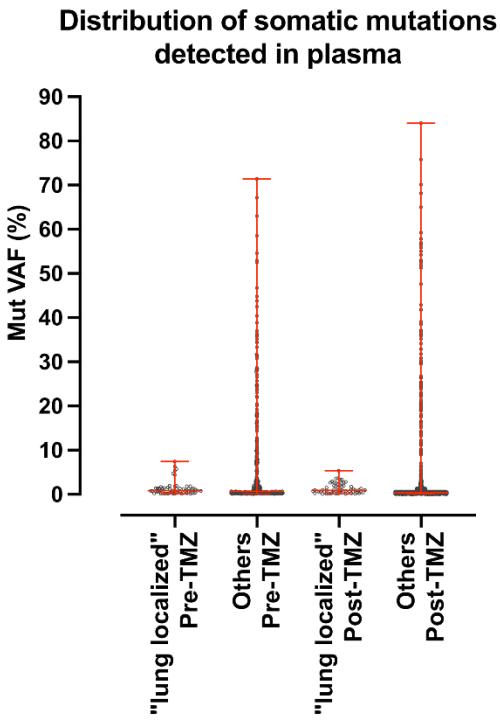
Mutational signatures in tumor biopsies obtained after TMZ priming**.**
A) Signature contribution analysis considering clonal mutations (fractional abundance ≥ 10%) for the heatmap generation and considering three different Signature reference databases (COSMIC ver. 2.0, ver. 3.0 and ver. 3.2). In five cases (AR02071, AR01032, AR01014, AR01069 and AR02064) the number of mutations was not sufficient to properly perform mutational analysis (cosine similarity lower than 0.9) and these five samples were excluded. **B)** Signature contribution analysis considering clonal and subclonal mutations (fractional abundance ≥ 1) for the heatmap generation, using three different signatures reference databases (COSMIC ver. 2.0, ver. 3.0 and ver. 3.2). Signature 11 related to alkylating agent was highlighted. *Sig.=Signature; TMZ = Temozolomide; vers. = version.*

Supplementary Figure S4

A



B



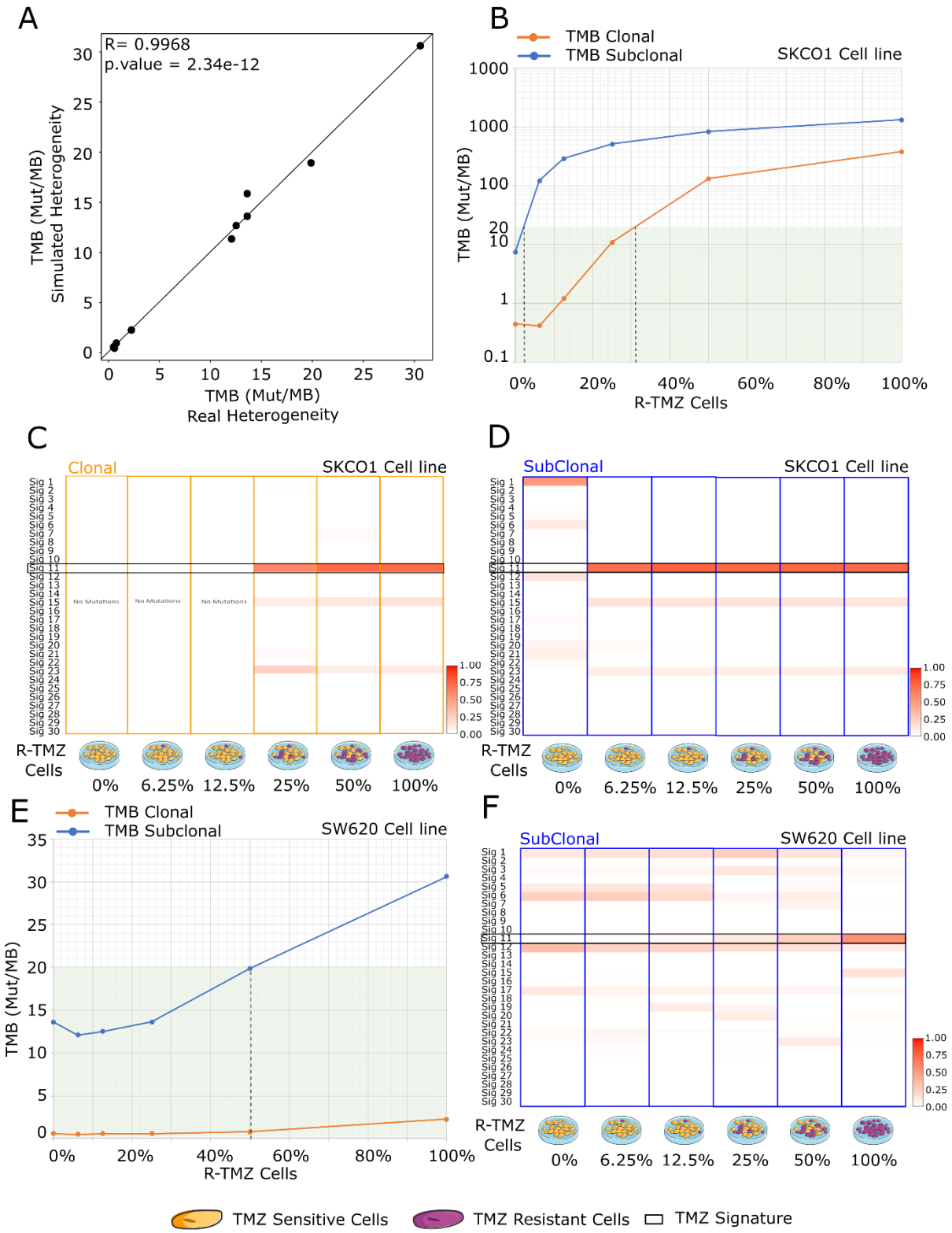
	"lung localized" Pre-TMZ	Others Pre-TMZ	"lung localized" Post-TMZ	Others Post-TMZ
Number of values	60	393	62	3531
Minimum	0.1400	0.06000	0.1000	0.01000
25% Percentile	0.3400	0.2900	0.3725	0.2300
Median	0.8550	0.6800	0.9250	0.3400
75% Percentile	1.400	6.765	1.988	0.6300
Maximum	7.420	71.42	5.350	84.04
Range	7.280	71.36	5.250	84.03

Supplementary Figure S4:

A) Frequency of alterations in the nucleotide context modified by TMZ signature 11 compared to all nucleotide contexts for *APC* and the MMR genes. Sum of frequency of all favored contexts in the Signature 11 were reported (Methods for details). **B)** Distribution of the somatic mutations identified in plasma collected before and after the TMZ treatment in

patients with lesions exclusively localized in lung (“Lung localized”) and in patients with lesions localized also or only in liver (“Others”). *Mut = Mutation; VAF = Variant Allelic Frequency; TMZ = Temozolomide; X is the average value*

Supplementary Figure S5:

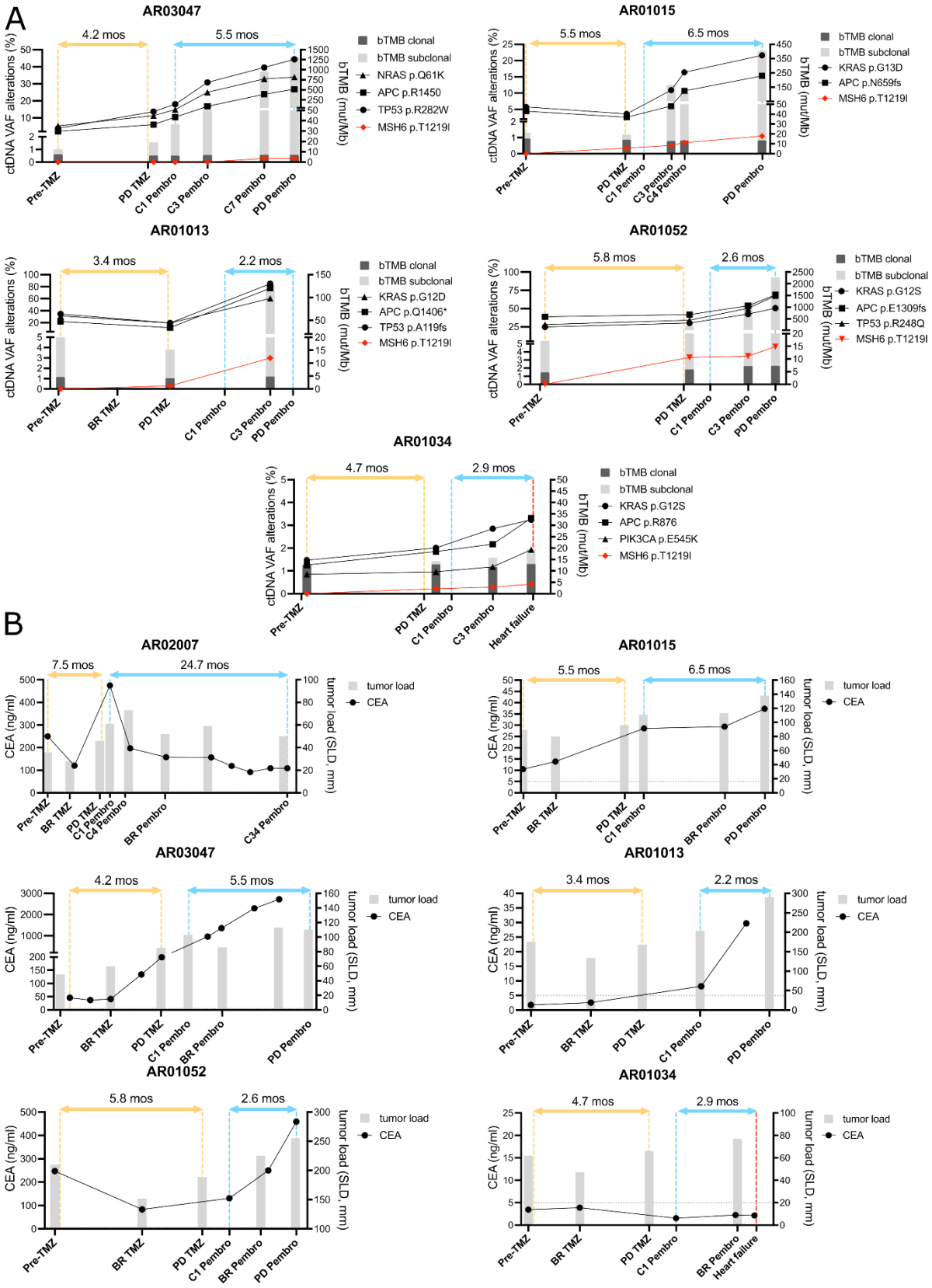


Supplementary Figure S5:

Impact of molecular heterogeneity on mutational signatures and TMB in CRC cell lines treated according to a TMZ schedule which parallels ARETHUSA therapy. **A):** Positive

Linear Correlation between TMB calculated using real and simulated data, Pearson's product-moment correlation is reported (R = 0.9968) with p-value = 2.34e-12. **B)** Increase of TMB observed by extending the fractional abundance of the TMZ resistant SKCO1 cells **C-D):** Effect of the heterogeneity on the Signature 11. By increasing the fractional abundance of TMZ resistant cells the contribution of signature 11 increased. The signature contribution was calculated considering only clonal mutation in panel **C** and considering all genetic alteration in panel **D** for the SKCO1 cell line. **E):** Effect of heterogeneity on TMB. Increase of Tumor Mutational Burden enlarging the fractional abundance of the TMZ resistant cells in SW620 cell line**. F):** Effect of the heterogeneity on the Signature 11 (related to Temozolomide). By increasing the fractional abundance of TMZ resistant cells the contribution of signature 11 increased. The signature contribution was calculated considering all genetic alteration for the SW620 cell line. *Sig.=Signature; TMZ = Temozolomide; TMB= Tumor Mutational Burden; R-TMZ= resistant to temozolomide treatment; Mut/MB = Mutations for megabases*

Supplementary Figure S6:

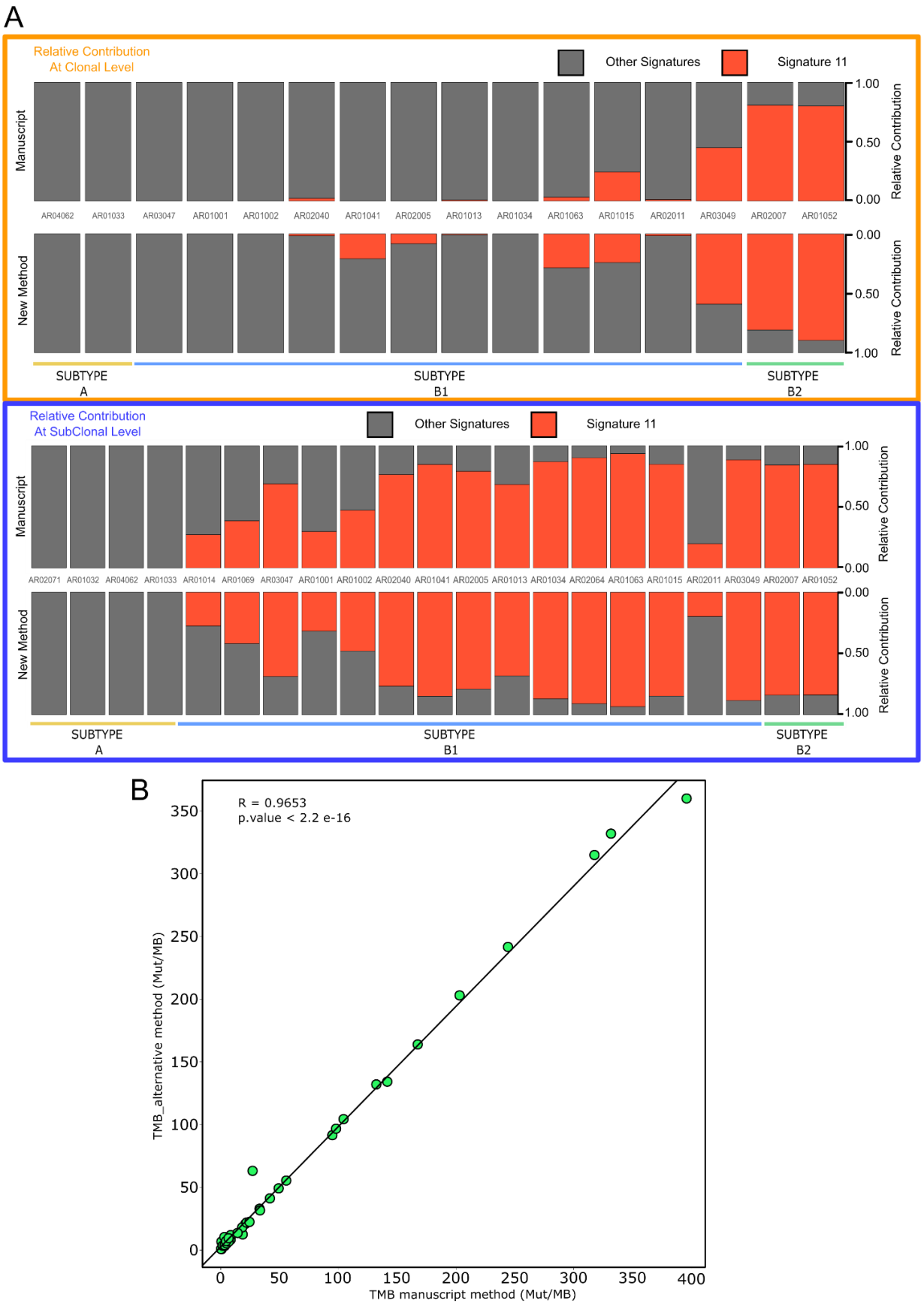


Supplementary Figure S6:

A: Graphs show longitudinal, liquid biopsy-based ctDNA monitoring of the patients AR03047, AR01015, AR01013, AR01052 and AR01034 during TMZ priming and

immunotherapy. Colored lines indicate clonal evolution of trunk/driver mutations. *KRAS, TP53; APC; NRAS PIK3CA* are shown in black and *MSH6* p.T1219I variant is in red. bTMB (clonal and subclonal) at each time-point is also reported (dark and light gray bar). **B:** Longitudinal monitoring of CEA levels (ng/ml) and tumor load (SLD, mm) in patients during TMZ and pembrolizumab treatment. *Mo= Months. ORR = Objective Response Rate; SD= Stable Disease; PD=Progression Disease.*

Supplementary Figure S7:



Supplementary Figure S7:

A: Relative contribution of Signature 11 at clonal and subclonal level as reported in the manuscript and using an independent method based on (89). In five cases (AR02071,

AR01032, AR01014, AR01069 and AR02064) the number of mutations was not sufficient to properly perform mutational analysis and these five samples were excluded (cosine similarity lower than 0.9). **B:** Scatter plot showing the Linear positive correlation between TMB calculated as reported in the manuscript and using an independent method based on (89). A linear positive correlation was found R=0.9653, p.value 2.2 e-16, Spearman's rank correlation.

Supplementary Table S1

Clinical feature		Evaluated Patients (n=21)	
Age (median, range)		62	28-84
Gender (n; %)	M	10	47,6
	F	11	52,4
Primary T location (n; %)	Rectum	4	19,0
	Colon	17	81,0
Sidedness (n; %)	Left	14	66,7
	Right	7	33,3
MSS/MMR status (n; %)	MMS/MMRp	21	100,0
MGMT status (median, range)	ICH score	15	(0-75)
	MB score	78,0%	(41,2-100,0)
RAS mutation (n; %)	G12	13	61,9
	G13	4	19,0
	other	4	19,0
Baseline Metastasis localization (n; %)	thoracic only	2	9,5
	abdomen only	0	0,0
	liver only	0	0,0
	liver+abdomen	3	14,3
	liver+lung	6	28,6
	diffuse	10	47,6

Sample_ID	READS	Depth (X)	Coverage 1X (%)	Coverage 10X (%)	Coverage 100X (%)
AR01001_NORMAL	137970480	398	97,93	97,60	96,74
AR01001_postTMZ	152192214	419	97,95	97,63	96,92
AR01001_preTMZ	104877109	308	97,87	97,54	96,04
AR01002_NORMAL	106327402	315	97,99	97,64	96,19
AR01002_postTMZ	112052167	327	98,02	97,69	96,15
AR01002_preTMZ	159382962	448	98,04	97,71	97,02
AR01013_NORMAL	102499935	307	98,05	97,69	95,57
AR01013_postTMZ	131734771	376	98,06	97,68	95,77
AR01013_preTMZ	120746214	236	97,96	97,54	91,88
AR01014_NORMAL	157500667	414	97,83	97,48	96,21
AR01014_postTMZ	133532979	368	97,83	97,46	95,79
AR01015_NORMAL	92516393	288	97,88	97,51	95,94
AR01015_postTMZ	96139405	298	97,89	97,52	96,16
AR01032_NORMAL	156046824	425	98,03	97,70	96,82
AR01032_postTMZ	149247970	445	98,02	97,68	96,94
AR01032_preTMZ	77448370	240	97,92	97,57	96,82
AR01033_NORMAL	158543286	442	98,10	97,74	96,96
AR01033_postTMZ	142797558	383	98,01	97,67	96,60
AR01034_NORMAL	118705353	341	98,03	97,67	96,42
AR01034_postTMZ	154048005	435	98,04	97,68	96,91

AR01041_NORMAL	109343403	376	98,47	98,06	97,11
AR01041_postTMZ	74171029	259	98,38	97,94	95,91
AR02005_NORMAL	127530660	370	97,90	97,57	96,69
AR02005_postTMZ_1	121508468	363	97,85	97,50	96,32
AR02005_postTMZ_2	127373174	372	97,84	97,49	96,08
AR02005_postTMZ_3	94134643	272	97,84	97,48	95,08
AR02007_NORMAL	95334938	291	97,82	97,46	95,84
AR02007_postTMZ	110226608	332	97,85	97,49	96,16
AR02011_NORMAL	182153134	503	98,07	97,75	97,05
AR02011_postTMZ	87167583	202	97,64	97,19	91,47
AR02040_NORMAL	128296851	432	98,53	98,11	97,31
AR02040_postTMZ	110294428	359	98,50	98,08	97,02
AR03047_NORMAL	123322335	414	98,59	98,21	97,32
AR03047_postTMZ	88348171	302	98,51	98,13	96,63
AR03049_NORMAL	111500850	364	98,68	98,24	97,08
AR03049_postTMZ	145085723	444	98,72	98,30	97,28
AR03049_preTMZ	136694718	433	98,65	98,25	97,29
AR01052_NORMAL	193787690	499	98,40	98,02	97,15
AR01052_postTMZ	183422134	457	98,33	97,94	96,83
AR04062_postTMZ	110640087	349	98,50	98,09	96,68
AR04062_NORMAL	132059504	417	98,56	98,16	97,16
AR01069_postTMZ	124020946	365	98,65	98,24	96,82
AR01069_NORMAL	140021771	412	98,66	98,27	97,14
AR02064_NORMAL	136445500	432	98,69	98,28	97,16
AR02064_postTMZ	122566050	395	98,58	98,18	97,08
AR01063_NORMAL	156283925	490	98,63	98,22	97,31
AR01063_postTMZ	110074395	346	98,50	98,08	96,75
AR02071_postTMZ	120976134	385	98,71	98,31	97,32
AR02071_NORMAL	130673258	420	98,72	98,33	97,45

Supplementary Table S1:

Sheet1: Main clinicopathological features of patients enrolled in the ARETHUSA trial between April 2019 and December 2021 who underwent tissue biopsy following progressive disease after TMZ priming. *MMRp=mismatch repair proficient; MSS=microsatellite stable; MGMT= O[6]-methylguanine-DNA methyltransferase ; IHC=immunohistochemistry; MB=methyl beaming; n = number; TMB=Tumor Mutational Burden; PD=progressive disease; mut=mutant; N.A = not available; M = male; F = female;*

Sheet2: Summary of the NGS parameters of tissue samples. The numbers of sequenced reads, median Depth (X) and median coverage at median 1X and 10X and 100X depth are reported for each sample. *preTMZ= before temozolomide treatment, postTMZ= after temozolomide treatment;*

Patient ID	Age	Gender	Primary T localization	*RAS* mutation	Subtype	Signature TMZ (subclonal, Score)	TMB (subclonal, Mut/MB)	TMZ		bTMB (Mut/MB)		*RAS* extended mut		Lesions localization (new lesions post-TMZ)				
								N cycles	BR	pre-TMZ	post-TMZ	pre-TMZ	post-TMZ	Lung	Liver	Abdomen	other	Lymph nodes
AR02071	67	M	left colon	KRAS G12D	A	0	19	2	PD	42,12	37,13	2,07	2,68	x			colon	abdominal
AR01032	28	M	rectum	KRAS G13D	A	0	22	1	PD	18,18	26,79	26,55	39,07	x		x		thoracic; abdominal
AR04062	58	F	left colon	KRAS A146T	A	0	25	2	PD	10.53	22,02	32,97	54,67	x	x (❖)	x		
AR01033	65	M	left colon	KRAS T50I	A	0	30	2	PD	29,69	26,79	39,16	39,88	x (2)	x (1)	x		(thoracic; abdominal)
AR01014	46	F	rectum	KRAS G12D	B1	0,31	14	2	PD	17,24	45,01	23,74	55,04	x	x (❖)	x	bone	abdominal
AR01069	51	M	left colon	KRAS G13D/K117N	B1	0,39	26	4	SD	6,73	10,53	0,39	3,18		x	x	(bone)	
AR03047	53	M	right colon	NRAS Q61K	B1	0,53	36	4	SD	12,44	19,14	5,35	11,33	x	x (❖)			
AR01001	58	F	left colon	KRAS G13D	B1	0,27	38	1	PD	49,76	46,92	52,96	57,81	x	x (3)			
AR01002	49	M	left colon	KRAS G12V	B1	0,3	45	4	SD	12,44	82,37	18,11	18,13	x	(2)	x		thoracic
AR02040	67	F	rectum	KRAS G12V	B1	0,72	61	2	PD	21,05	60,29	44,83	52,88	x	x	x		
AR01041	46	F	right colon	KRAS A59E	B1	0,77	101	2	PD	26,81	62,24	54,57	59,24		x (❖)	x		abdominal
AR02005	67	F	right colon	KRAS G12D	B1	0,72	106	6	SD	NA	188,97	NA	32,86		x	x		
AR01013	69	M	rectum	KRAS G12D	B1	0,63	113	4	SD	20,11	15,31	31,33	19,37	x	x (1)			abdominal; cervical (1)
AR01034	82	M	right colon	KRAS G12S	B1	0,84	137	5	SD	12,69	14,25	1,47	2,01	x	x			
AR02064	62	M	right colon	KRAS G12D	B1	0,77	145	5	SD	4,78	95,26	1,88	4,58	x	x	x		
AR01063	68	F	right colon	KRAS G12D	B1	0,87	176	6	SD	24,88	67,03	27,1	35,19		x		spleen	
AR01015	50	F	left colon	KRAS G13D	B1	0,84	210	6	SD	21,05	19,14	5,74	3,62	x				
AR02011	74	M	left colon	KRAS G12D	B1	0,23	249	6	SD	17,88	18,68	1,41	2,68	x				
AR03049	63	M	left colon	KRAS G12D	B1	0,86	324	4	SD	8,61	NA	10,36	NA	x	x			
AR02007	84	F	left colon	KRAS G12V	B2	0,85	381	6+2	SD	25,84	2276,55	7,25	19,1	x	x (❖)			
AR01052	56	F	left colon	KRAS G12S	B2	0,84	423	6	SD	17,24	195,63	24,89	30,14	x (2)	x (2)		bone	thoracic

Supplementary Table S2:

For each patient, Signature stratification, bTMB, signature score of post TMZ biopsy, best clinical response to TMZ treatment with corresponding genetic localization of the metastases are reported. Patients with preferentially lung localization of the metastases and low shedding of ctDNA in the blood are highlighted in red. *M= male; F= Female; PD = patient with progression disease as best clinical response. SD= patient with stable disease as best clinical response; preTMZ= analysis of tumor before TMZ priming; Post-TMZ= analysis of tumor after TMZ priming; ❖ = Multiple nodules; bTMB = blood Tumor Mutational Burden; VAF = Variant Allelic Frequency*

Supplementary Table S3:

Genetic TMZ effect	Patient_ID	TMZ_Signature (score)	MSH6 alterations (post TMZ) in tissue	MSH6 alterations (pre TMZ) in blood	MSH6 alterations (post TMZ) in blood
SUBTYPE A: no genetic TMZ effect	**AR02071**	**0**	-	-	-
SUBTYPE A: no genetic TMZ effect	**AR01032**	**0**	-	-	-
SUBTYPE A: no genetic TMZ effect	**AR04062**	**0**	-	-	-
SUBTYPE A: no genetic TMZ effect	**AR01033**	**0**	-	-	-
SUBTYPE B1: subclonal effect	**AR01014**	**0,31**	-	-	p.T1219I
SUBTYPE B1: subclonal effect	**AR01069**	**0,39**	p.K301fs*12	-	-
SUBTYPE B1: subclonal effect	**AR03047**	**0,53**	-	-	p.T1219I[1]
SUBTYPE B1: subclonal effect	**AR01001**	**0,27**	-	-	p.T1219I
SUBTYPE B1: subclonal effect	**AR01002**	**0,30**	-	-	p.T1219I
SUBTYPE B1: subclonal effect	**AR02040**	**0,72**	-	-	p.T1219I
SUBTYPE B1: subclonal effect	**AR01041**	**0,77**	-	-	p.T1219I
SUBTYPE B1: subclonal effect	**AR02005**	**0,72**	p.T1219I;p.G557D	NA	p.T1219I;p.G557D; p.Q626*; p.W456*; p.W413*; p.G520D; p.G1157D; p.T1008I
SUBTYPE B1: subclonal effect	**AR01013**	**0,63**	-	-	p.T1219I
SUBTYPE B1: subclonal effect	**AR01034**	**0,84**	p.T1219I;p.G557D; p.V777*; p.W50*	-	p.T1219I
SUBTYPE B1: subclonal effect	**AR02064**	**0,77**	-	-	Q1122*; p.T1219I
SUBTYPE B1: subclonal effect	**AR01063**	**0,87**	p.W1047*	-	p.T1219I;p.S1188N; p.Q485*
SUBTYPE B1: subclonal effect	**AR01015**	**0,84**	p.T1219I	-	p.T1219I
SUBTYPE B1: subclonal effect	**AR02011**	**0,23**	p.T1219I	-	p.T1219I
SUBTYPE B1: subclonal effect	**AR03049**	**0,86**	p.T1219I	-	NA
SUBTYPE B2: clonal effect	**AR02007**	**0,85**	p.G1139S	-	p.G1139S; p.T1219I; p.Q1122*; p.W142*; p.Q1314*; p.W372*; p.Q244*
SUBTYPE B2: clonal effect	**AR01052**	**0,84**	-	-	p.T1219I;p.T1008I

Supplementary Table S3:
Molecular features of ARETHUSA patients: For each patient, the molecular effect of TMZ, the relative contribution of Signature 11 in tissue samples after TMZ and occurrence of *MSH6* mutations identified in the blood (red color) before and after the TMZ treatment are listed. Variants of unknown significance were filtered and only mutations likely to inactivate MMR system are listed. *Sig.=Signature; TMZ = Temozolomide; Subtype A= patients with no molecular evidence of Temozolomide treatment. Subtype B1= patients with subclonal molecular evidence of Temozolomide treatment. Subtype B2= patients with clonal molecular evidence of Temozolomide treatment. NA = Not Available. [1]identified during Pembrolizumab longitudinal monitoring (at cycle 7).*