

Figure S1: **Heatmap clustering of gastric cancer sample signature exposure.** Mutational signatures called by signeR are arranged in rows, and samples are arranged in columns.

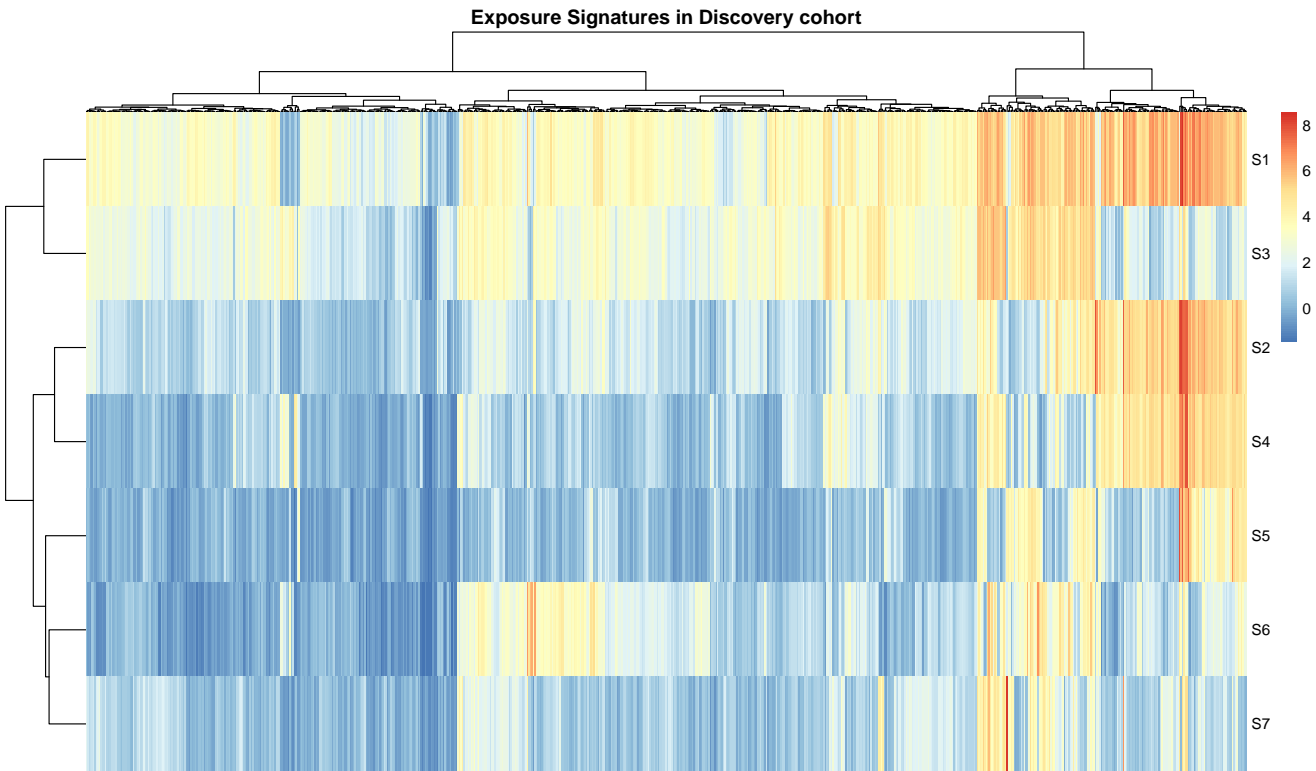


Figure S2: Forest plot of the hazard ratios for overall survival from univariate Cox model analysis. TMB, tumor mutational burden; MSI, microsatellite instability.

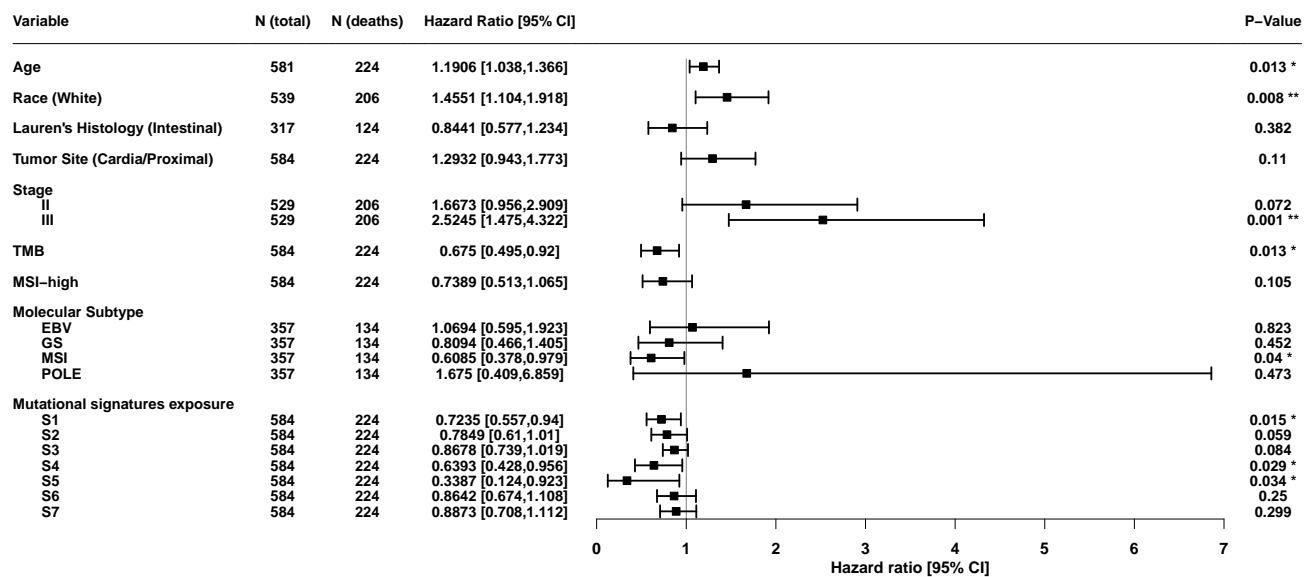
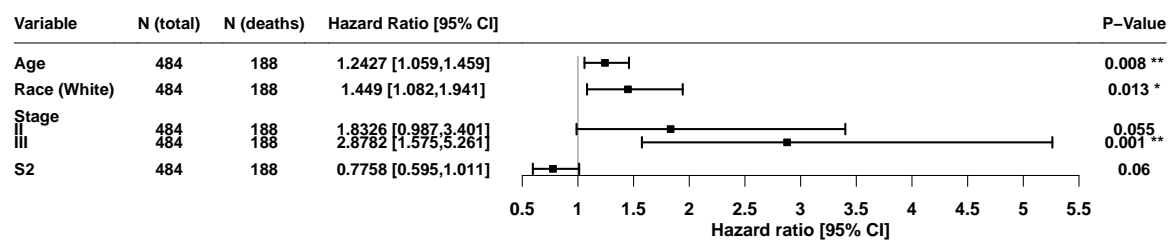


Figure S3: Forest plot of hazard ratios for overall survival from the multivariate Cox model for (a) mutational signature (S) 2 and (b) S5.

(a) Signature S2



(b) Signature S5

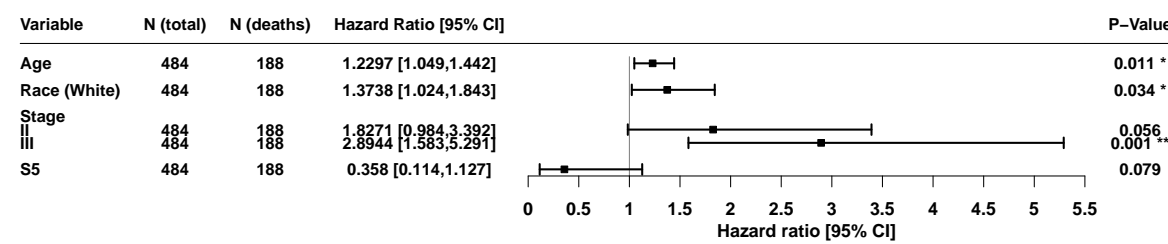


Figure S4: Calibrated plots of 2-year survival (multiple Cox regression models) for (a) mutational signature (S)2, (b) S4, (c) S5, (d) high microsatellite instability, and (e) tumor mutational burden.

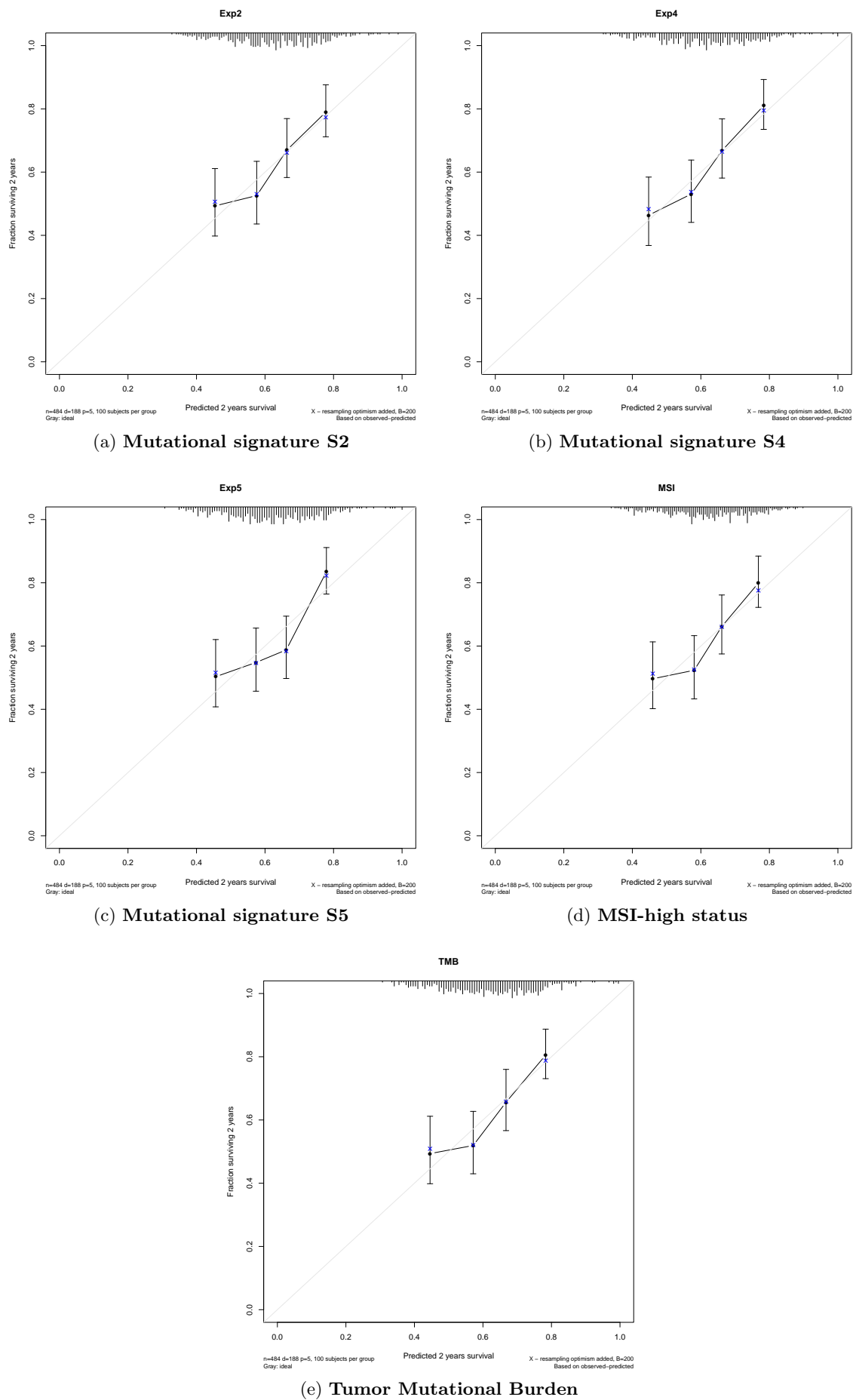


Figure S5: Five-year overall survival for $S4^{high}$ and $S4^{low}$ groups in the public cohort.

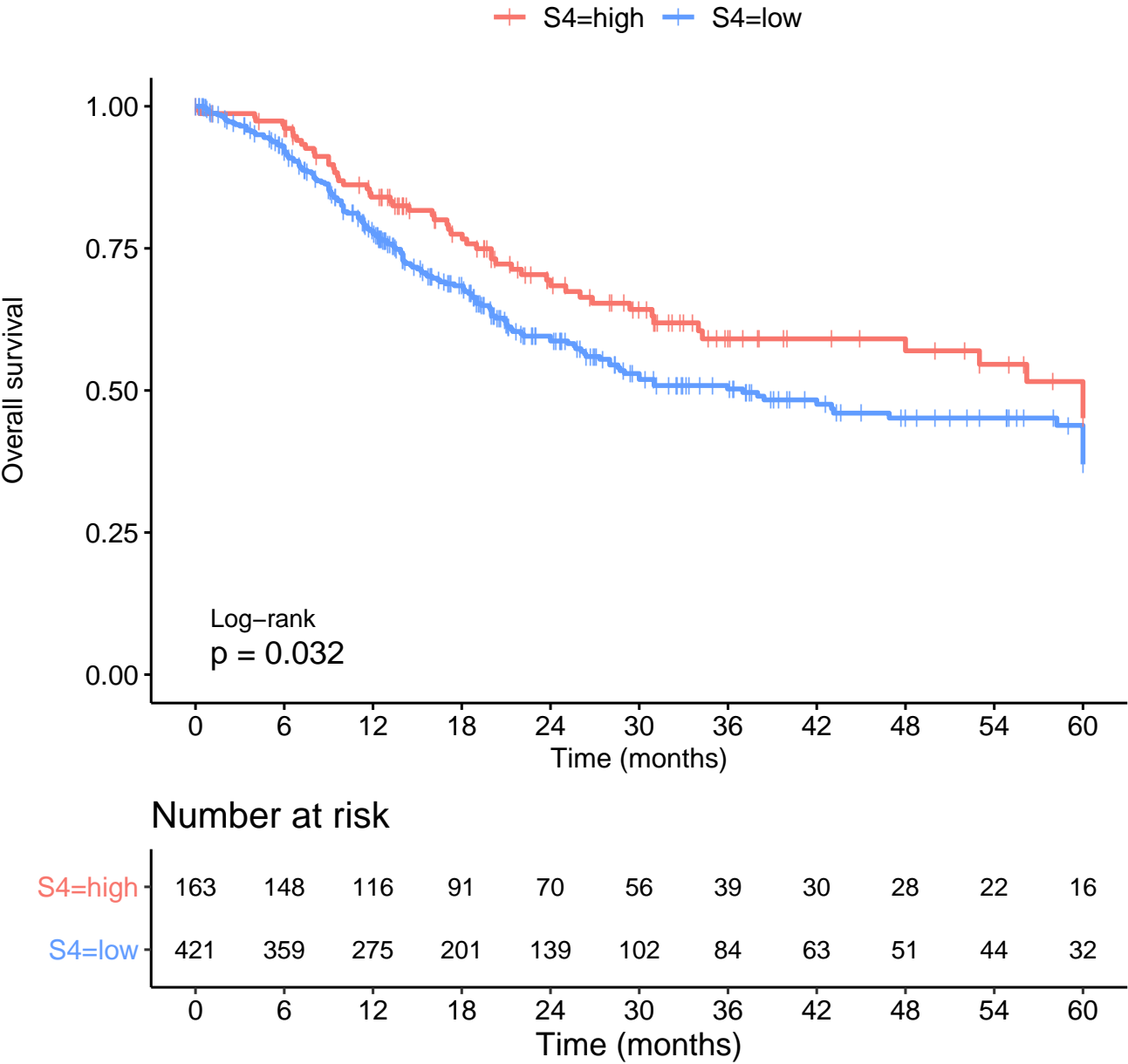


Figure S6: Five-year overall survival for $S4^{high}$ and $S4^{low}$ groups in the validation cohort.

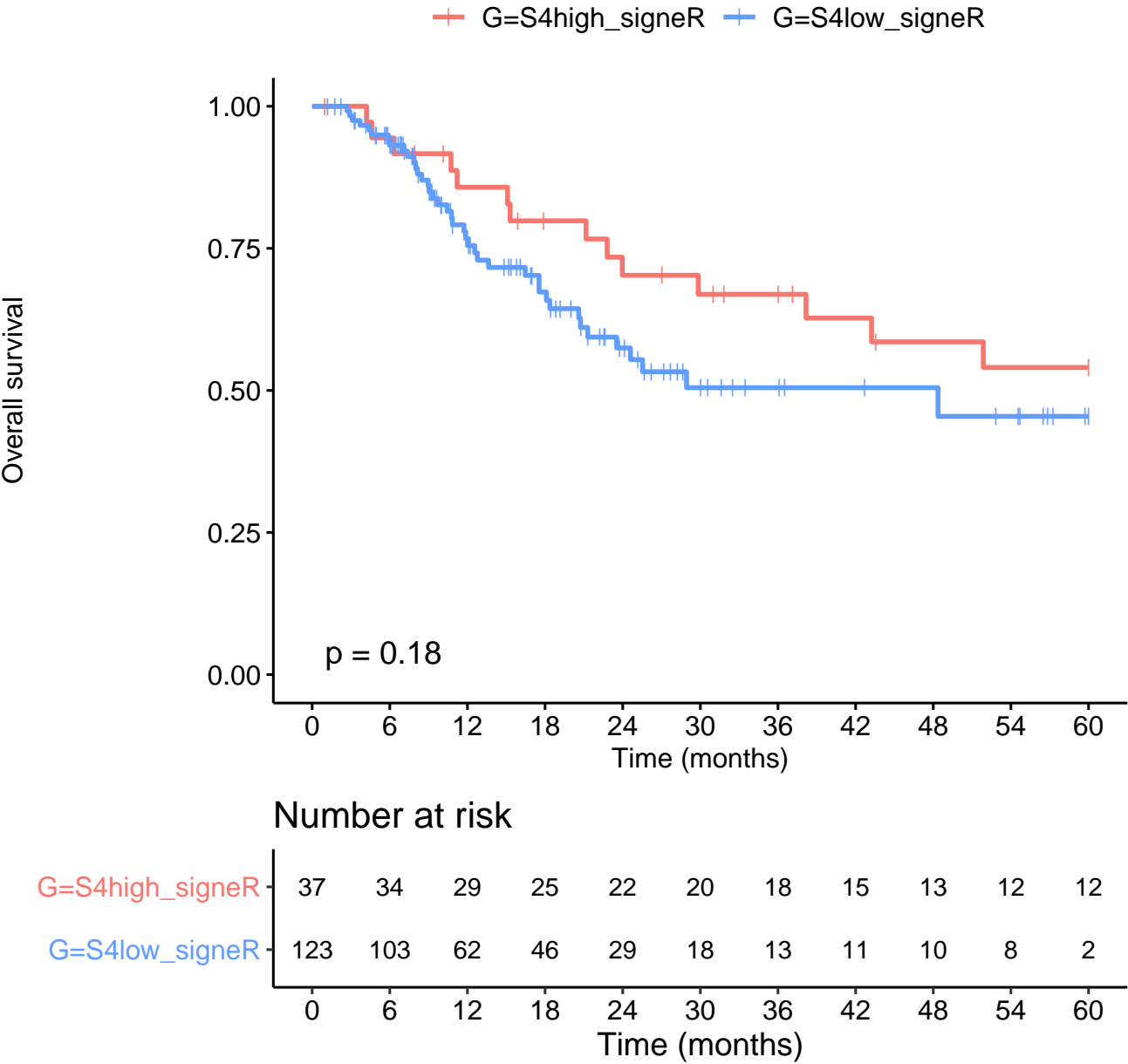


Figure S7: Scatter plots of Spearman's correlation between MMR gene expression levels - FPKM (log-scaled) and methylation loads. Gene expression levels were plotted against average β values for the promoter-related CpG islands of MMR genes.

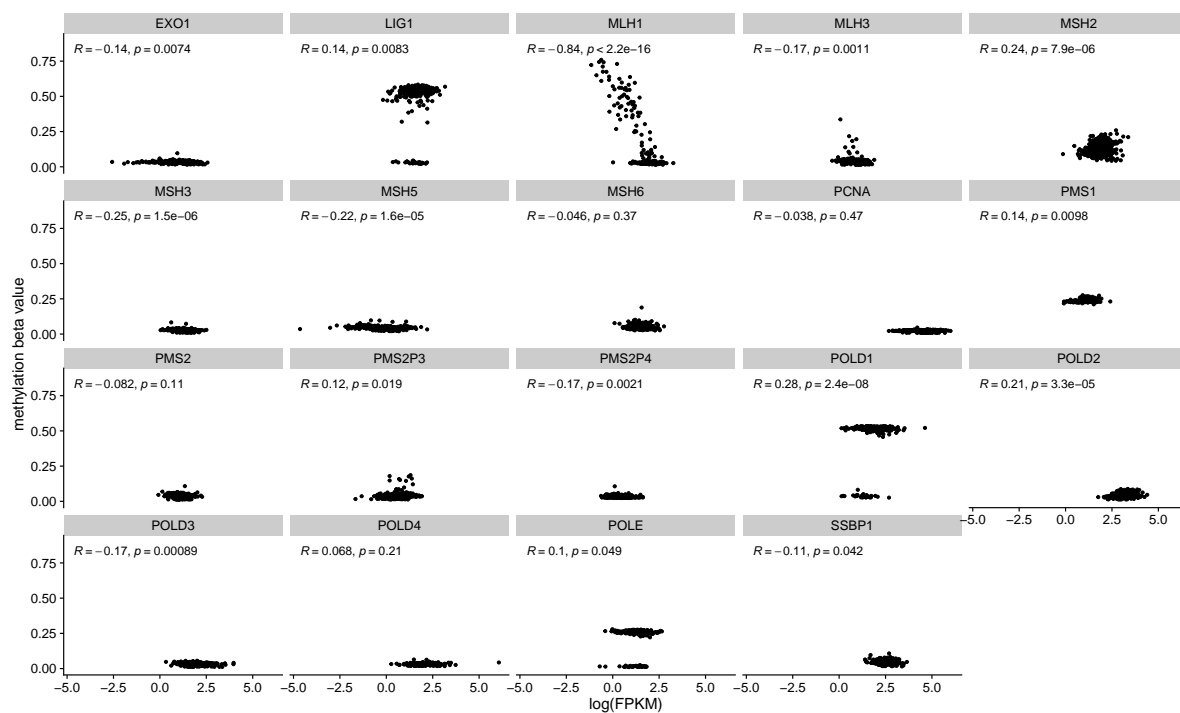


Figure S8: Barplot showing the burdens of previously identified mutational signatures previously identified in the public gastric cancer cohort; isogenic wild-type cells are compared with two subclones with mutations in the *MLH1* gene induced by CRISP-Cas9 assay.

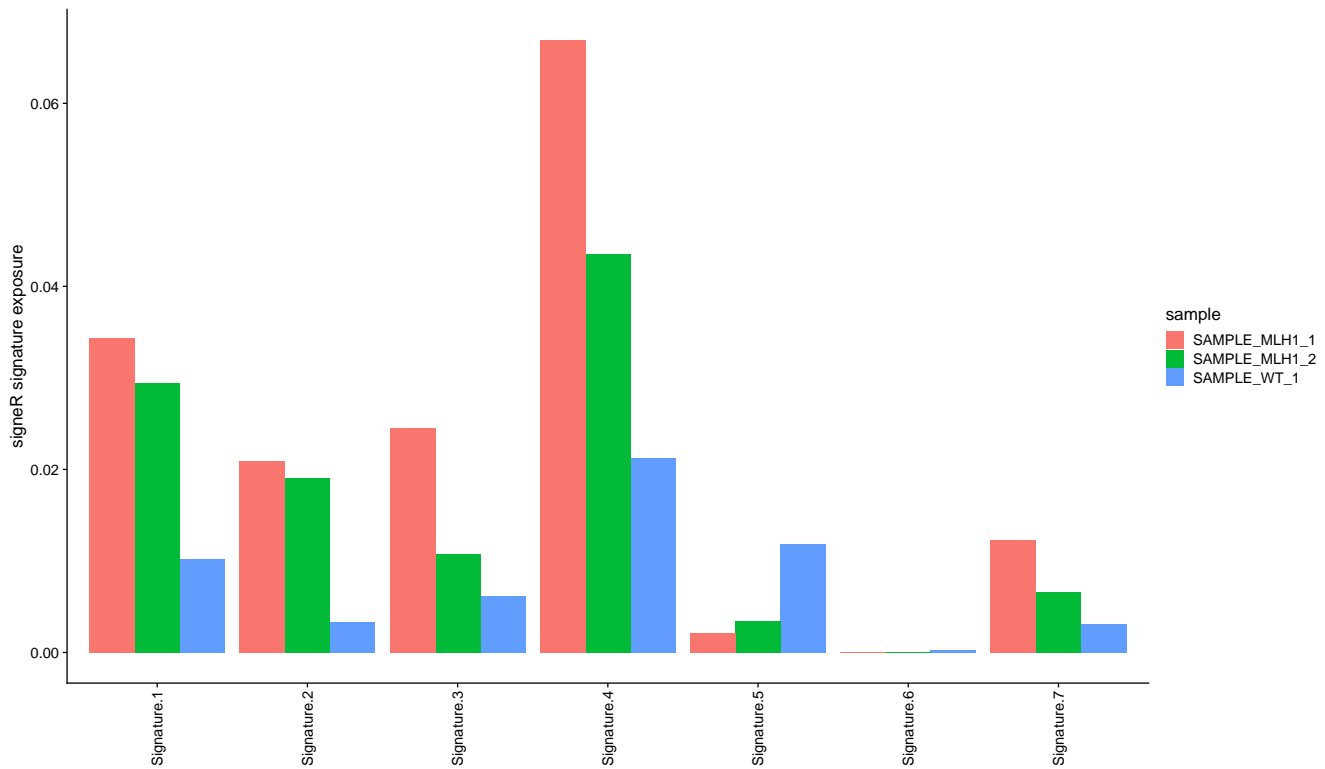
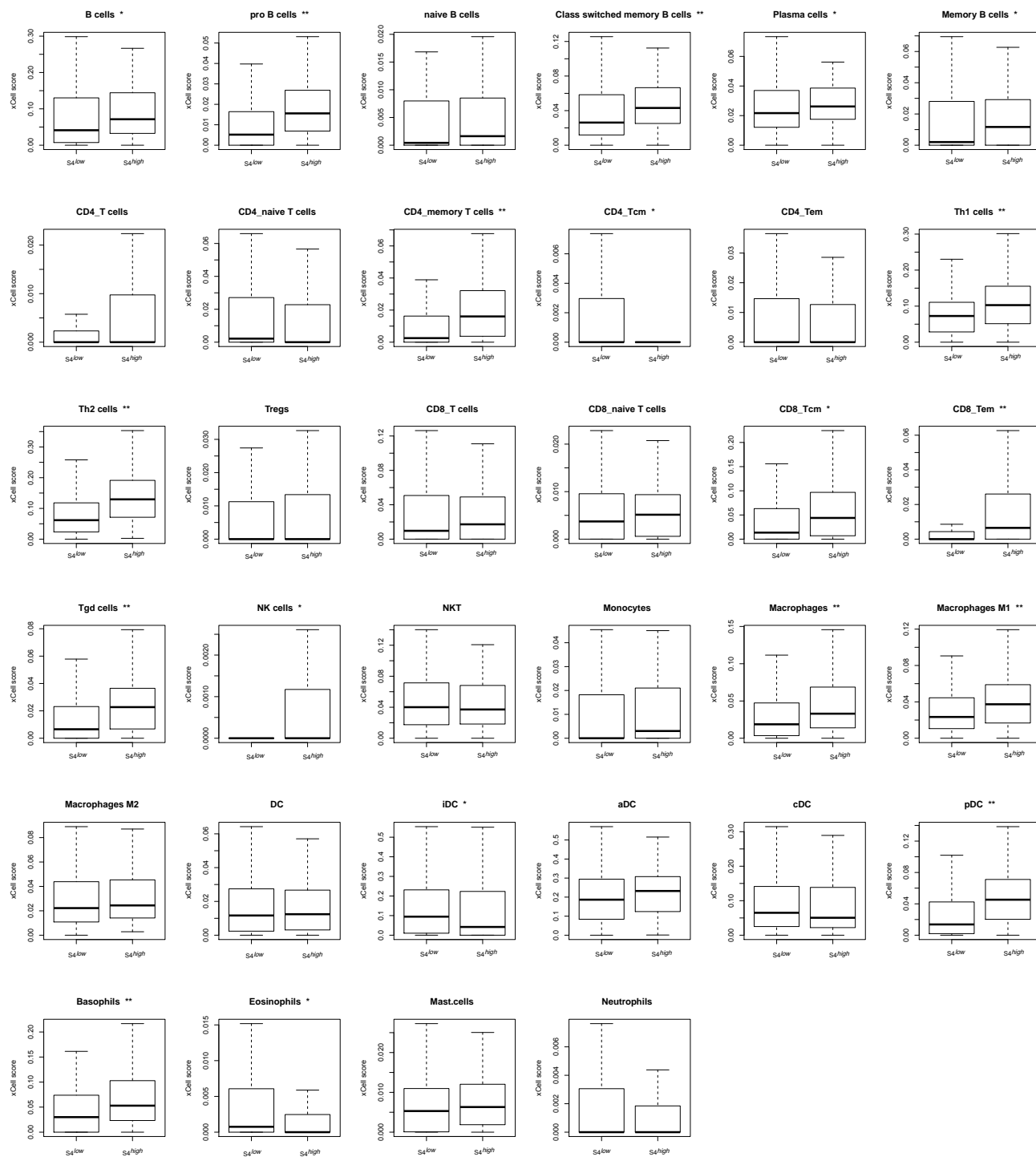
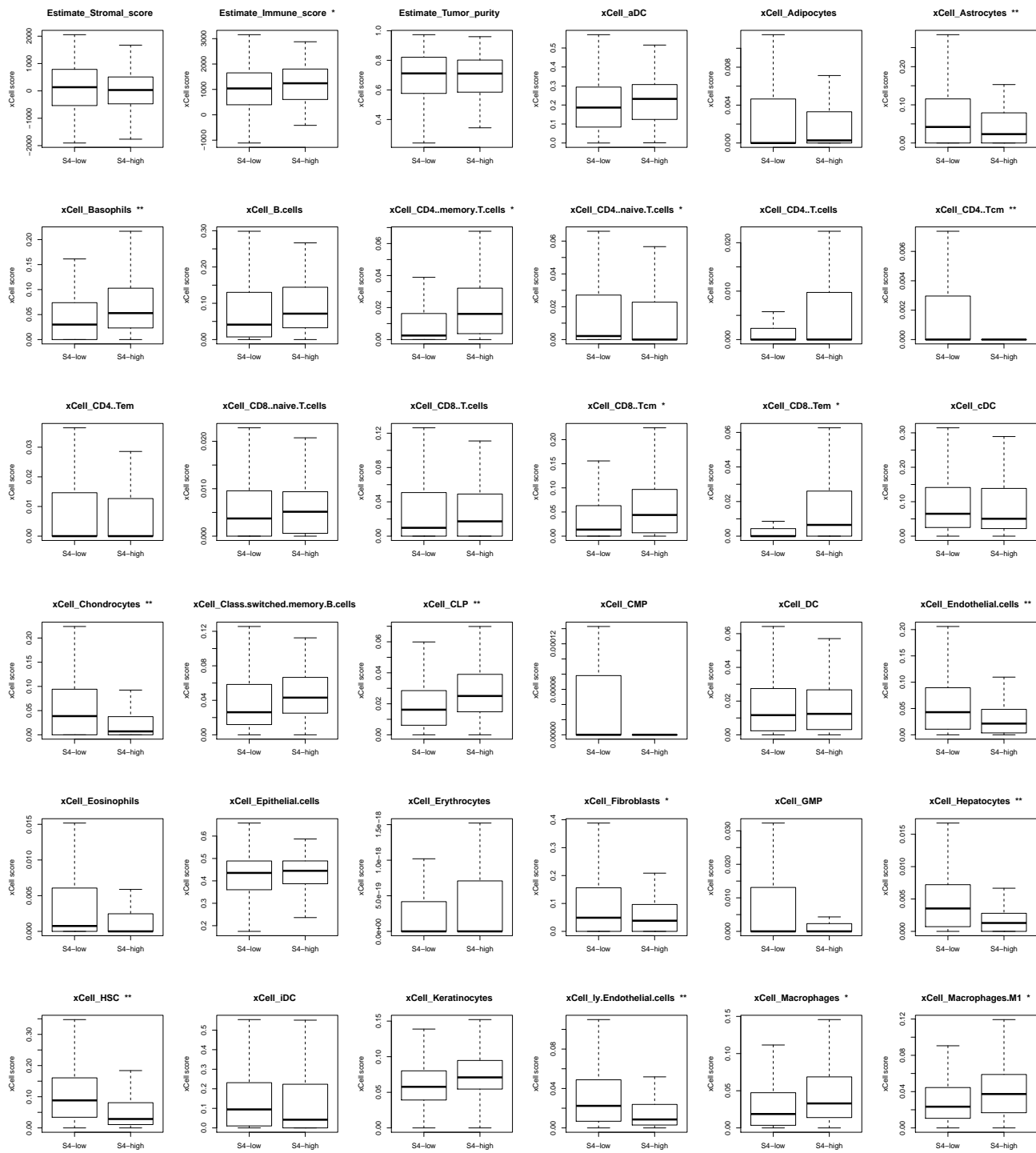


Figure S9: Boxplots of immune and stroma cells scores, estimated by xCell [?], in the mutational signature ($S4^{high}$ and $S4^{low}$ groups. * $p < 0.05$, ** $p < 0.01$, t test.





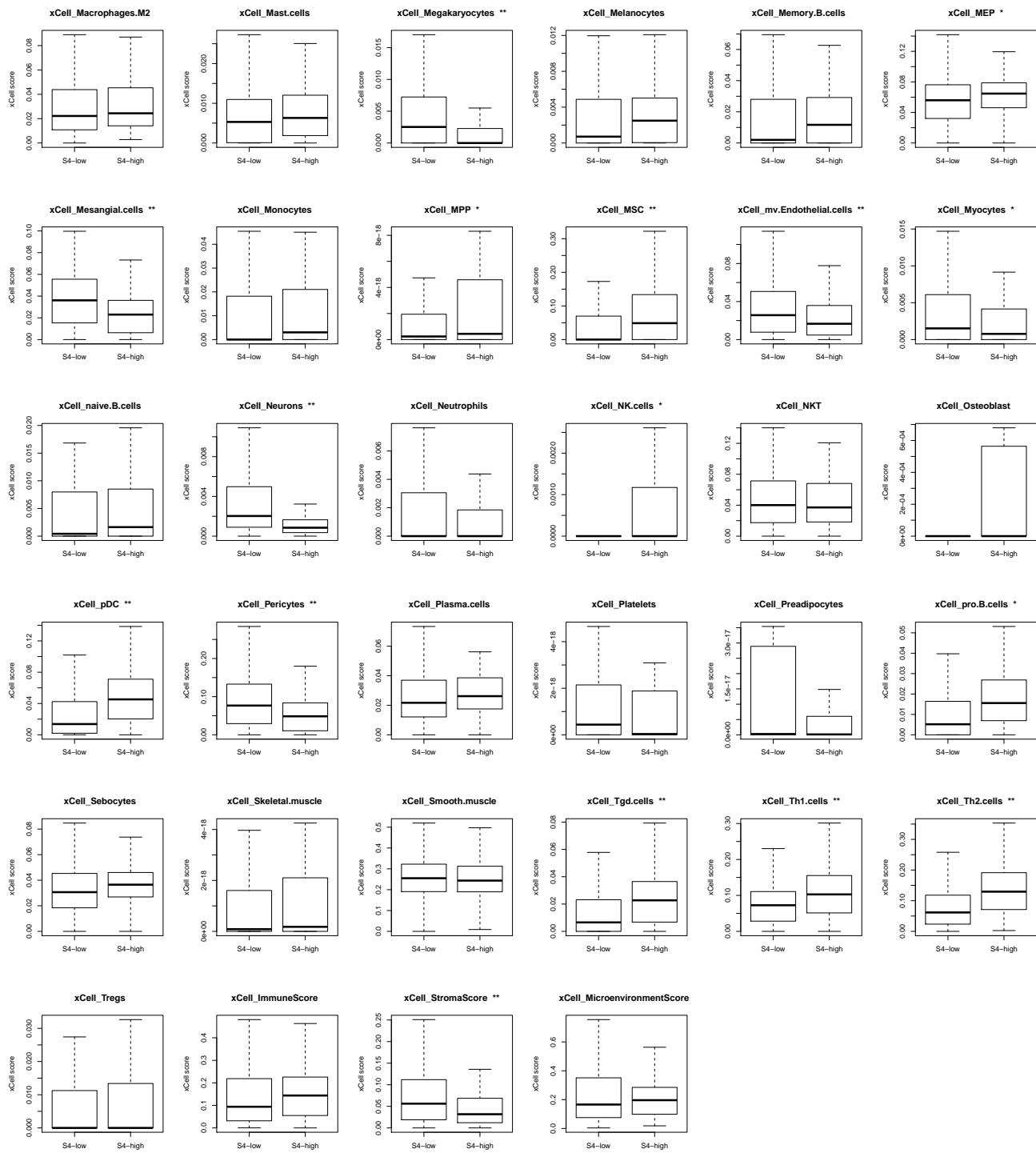


Figure S10: **Boxplots showing the normalized gene expression of immune-related genes.** Immune gene markers as (a) inflammatory/cytotoxic, (b) suppressor/exhausted, (c) costimulatory/antigen presentation, and (d) other natural killer cell receptors and monocyte/macrophage markers. Outlying values were removed.* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, Mann-Whitney U test.

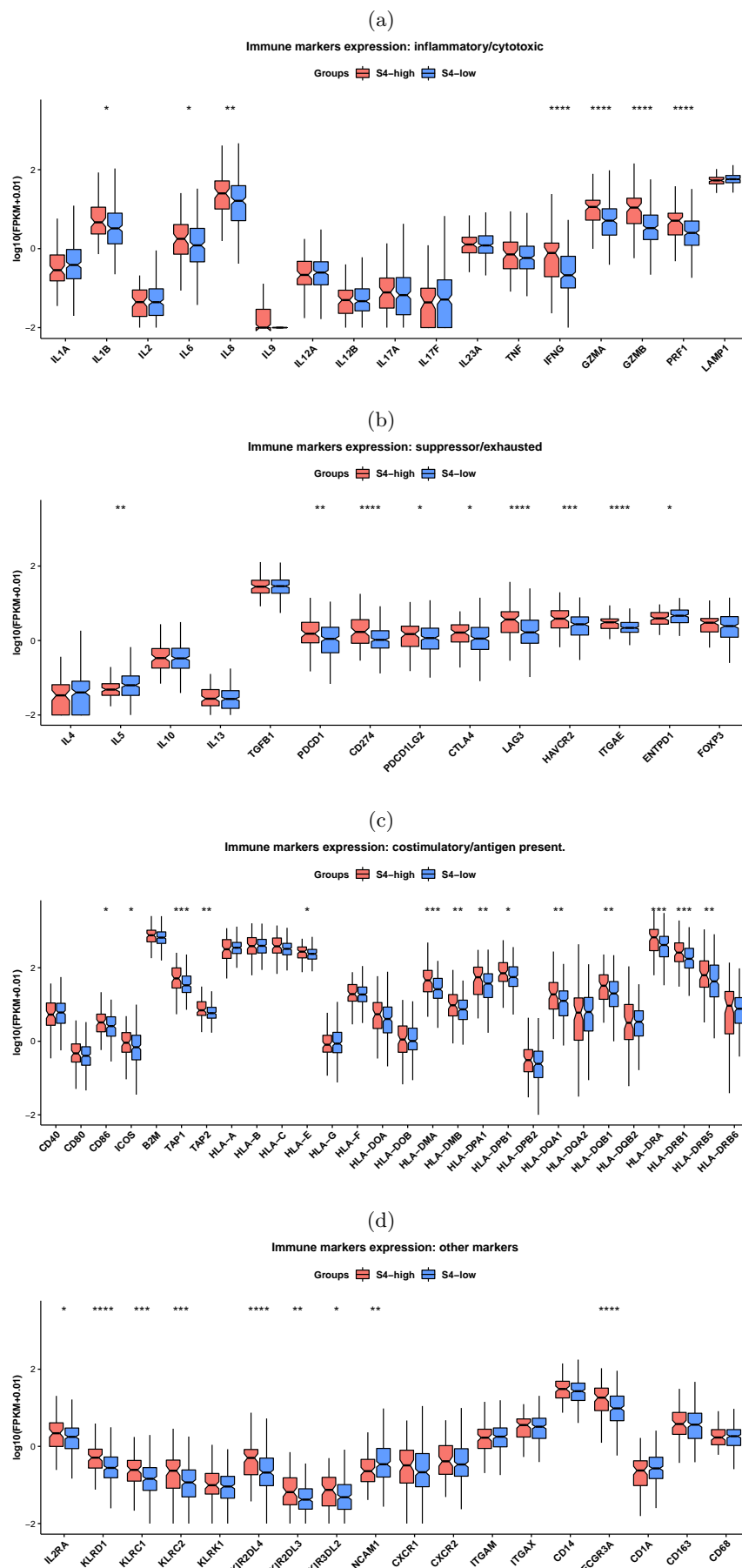


Table S1: Metadata on sample clinical, molecular and immune features.

TableS1_Clinical-Molecular-features.xlsx.

Table S2: **Significantly mutated genes in the mutational signature (S)4^{high} and S4^{low} groups and associated pathways.** (a) Significantly mutated genes, related pathways and frequency in the S4^{high} group. (b) Description of pathways found for the S4^{high} gene set. (c) Significantly mutated genes, related pathways and frequency in the S4^{low} group. (d) Description of pathways found for the S4^{low} gene set.

TableS2_SignificantlyMutatedGenes-Pathways.xlsx.

Table S3: Clinicopathological features of gastric cancer samples by cohort.

	TCGA(<i>n</i> = 439)		cBioPortal(<i>n</i> = 226)		ICGC(<i>n</i> = 122)	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Age (mean ± SD)	65.75 ± 10.69		63.1 ± 13.73		60.30 ± 10.74	
Gender	439	100	206	91	122	100
Female	158	36	79	38	37	30
Male	281	64	127	62	85	70
Race	378	86	226	100	122	100
White	275	73	0	0	0	0
Black	13	3	0	0	0	0
Asian	89	24	226	100	122	100
Other	1	0.26	0	0	0	0
Anatomic Site	422	96	205	91	0	0
Cardia/Proximal	107	25	61	30	0	0
Fundus/Body	151	36	61	30	0	0
Antrum/Distal	159	38	83	40	0	0
Other	5	1	0	0	0	0
Histology Lauren	261	59	206	91	0	0
Diffuse	73	28	77	37	0	0
Intestinal	188	72	113	55	0	0
Mixed	0	0	16	8	0	0
Stage T	430	98	156	69	117	96
T1 – T2	115	27	42	27	24	20
T3 – T4	315	73	114	73	93	80
Stage N	421	96	156	69	117	96
N0	131	31	20	13	22	19
N+	290	69	136	87	95	81
Stage M	439	100	156	69	112	92
M0	387	88	133	85	103	92
M1	30	7	23	15	9	8
MX	22	5	0	0	0	0
Pathological Stage	413	94	205	90	97	79
I	61	15	15	7	97	79
II	165	40	38	18	17	17
III	157	38	72	35	60	62
IV	30	7	80	39	11	11
Molecular Subtype	383	87	20	9	0	0
CIN	223	58	0	0	0	0
GS	50	13	0	0	0	0
EBV	30	8	8	40	0	0
MSI	73	9	12	60	0	0
POLE	7	2	0	0	0	0
MSIseq Status	439	100	226	100	122	100
MSI – high	87	20	45	20	28	23
nonMSI – high	352	80	181	80	94	77
Immune Subtype	388	88	0	0	0	0
C1	128	33	0	0	0	0
C2	209	54	0	0	0	0
C3	35	9	0	0	0	0
C4	9	2	0	0	0	0
C6	7	2	0	0	0	0

Table S4: Clinicopathological features of gastric cancer samples according to mutational signature S4 DNA mismatch repair deficiency in the validation cohort

	<i>All</i> (<i>n</i> = 170)		<i>S4^{low}</i> (<i>n</i> = 127)		<i>S4^{high}</i> (<i>n</i> = 43)		<i>Pvalue</i>
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	
Age (mean ± SD)	61.06 ± 13.33		60.182 ± 13.24		63.65 ± 13.39		0.0726
Gender	170	100	127	100	43	100	
<i>Female</i>	64	38	45	35	19	44	0.3999
<i>Male</i>	106	62	82	65	24	56	
Race	154	91	120	94	34	79	
<i>White</i>	91	59	72	60	19	56	0.7449
<i>Black</i>	11	7	7	6	4	12	
<i>Asian</i>	20	13	16	13	4	12	
<i>Brown</i>	31	20	24	20	7	21	
<i>Other</i>	1	1	1	1	0	0	
Anatomic Site	170	100	127	100	43	100	
<i>Cardia/GEJ/Distalesophagus</i>	57	34	47	37	10	23	0.0396
<i>Anastomosicmouth</i>	3	2	3	2	0	0	
<i>Body</i>	62	36	48	38	14	33	
<i>Antrum</i>	44	26	28	22	16	37	
<i>Linitis</i>	4	2	1	1	3	7	
Histology Lauren	152	89	111	87	41	95	
<i>Diffuse</i>	74	49	58	52	16	39	0.2246
<i>Intestinal</i>	59	39	39	35	20	49	
<i>Mixed</i>	19	13	14	13	5	12	
Stage T	103	89	67	53	36	84	
<i>T1 – T2</i>	44	43	31	46	13	36	0.5326
<i>T3 – T4</i>	59	57	36	54	23	64	
Stage N	102	60	75	59	37	86	
<i>N0</i>	43	42	39	52	14	38	0.226
<i>N+</i>	59	57	36	48	23	62	
Stage M	111	65	74	58	37	86	
<i>M0</i>	27	24	4	5	23	62	<0.0001
<i>M1</i>	2	2	2	3	0	0	
<i>MX</i>	82	74	68	92	14	38	
Pathological Stage	105	62	68	53	37	86	
<i>I</i>	31	30	24	35	7	19	0.2194
<i>II</i>	35	33	23	34	12	32	
<i>III</i>	31	30	17	25	14	38	
<i>IV</i>	8	8	4	6	4	11	
MSIseq Status	170	100	127	100	43	100	
<i>MSI – high</i>	38	22	2	2	36	84	<0.0001
<i>nonMSI – high</i>	132	78	125	98	7	16	