



*Research article*

**Elevated expression of the MYB proto-oncogene like 2 (MYBL2)-encoding gene as a prognostic and predictive biomarker in human cancers**

**Zekun Xin<sup>1,†</sup>, Yang Li<sup>2,†</sup>, Lingyin Meng<sup>3,†</sup>, Lijun Dong<sup>1</sup>, Jing Ren<sup>2,\*</sup> and Jianlong Men<sup>2,\*</sup>**

<sup>1</sup> Urology Surgery, Hebei Petro China Central Hospital, Langfang 065000, China

<sup>2</sup> Precision Medicine Center, Tianjin Medical University General Hospital, Tianjin 300000, China

<sup>3</sup> Central Laboratory, Linyi People's Hospital, Linyi 276000, China

<sup>a</sup> Zekun Xin, Yang Li and Lingyin Meng contributed equally to this work.

\* **Correspondence:** Email: renjing\_tmu@163.com, menjianlong@163.com.

† These authors contributed equally to this work.

**Supplementary**

---

**Table S1.** Annotation TCGA cancer type.

---

TCGA	Detail
ACC	Adrenocortical carcinoma
BLCA	Bladder urothelial carcinoma
BRCA	Breast invasive carcinoma
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma
CHOL	Cholangio carcinoma
COAD	Colon adenocarcinoma
DLBC	Lymphoid neoplasm diffuse large B-cell lymphoma
ESCA	Esophageal carcinoma
GBM	Glioblastoma multiforme
HNSC	Head and neck squamous cell carcinoma
KICH	Kidney chromophobe
KIRC	Kidney renal clear cell carcinoma
KIRP	Kidney renal papillary cell carcinoma
LAML	Acute myeloid leukemia
LGG	Brain lower grade glioma
LIHC	Liver hepatocellular carcinoma
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
MESO	Mesothelioma
OV	Ovarian serous cystadenocarcinoma
PAAD	Pancreatic adenocarcinoma
PCPG	Pheochromocytoma and paraganglioma
PRAD	Prostate adenocarcinoma
READ	Rectum adenocarcinoma
SARC	Sarcoma
SKCM	Skin cutaneous melanoma
STAD	Stomach adenocarcinoma
TGCT	Testicular germ cell tumors
THCA	Thyroid carcinoma
THYM	Thymoma
UCEC	Uterine corpus endometrial carcinoma
UCS	Uterine carcinosarcoma
UVM	Uveal melanoma

---

**Table S2.** LogFC between TCGA and GTEx samples.

Primary Tissue	Deletion – Frameshift	Insertion - Frameshift	Substitution - coding silent	Substitution - Missense	Substitution - Nonsense	gain	loss	Unknown	Total samples	percentage mutation( %)
Adrenal gland	0	0	0	0	0	2	0	0	627	0
Autonomic ganglia	0	0	0	0	0	9	2	0	1135	0
Biliary tract	0	1	2	3	0	0	0	1	549	1.093
Bone	0	0	1	0	0	0	0	0	569	0.176
Breast	0	0	1	8	0	97	43	0	2591	0.347
Central nervous system	0	0	1	1	0	0	0	1	2457	0.081
Cervix	0	0	0	1	0	4	0	0	329	0.304
Endometrium	0	0	2	10	0	7	0	0	715	1.678
Eye	0	0	0	0	0	1	0	0	119	0
Fallopian tube	0	0	0	0	0	0	0	0	2	0
Gastrointestinal tract (site indeterminate)	0	0	0	0	0	0	0	0	1	0
Genital tract	0	0	0	0	0	0	0	0	68	0
Haematopoietic and lymphoid	0	0	7	4	0	0	1	1	4589	0.24
Kidney	0	0	3	7	1	12	6	0	2373	0.464
Large intestine	0	2	18	38	0	10 1	4	1	2358	2.46
Liver	0	0	2	11	2	3	0	3	2205	0.68
Lung	0	0	7	22	1	13	0	2	2698	1.112
Meninges	0	0	0	0	0	0	0	0	65	0
NS	1	0	2	4	1	0	0	0	133	6.015
Oesophagus	0	0	0	1	0	68	20	1	1665	0.06
Ovary	0	0	0	0	0	85	27	1	950	0
Pancreas	0	0	0	0	0	35 1	50	0	1872	0

Primary Tissue	Deletion - Frameshift	Insertion - Frameshift	Substitution - coding silent	Substitution - Missense	Substitution - Nonsense	gain	loss	Unknown	Total samples	percentage mutation( %)
Parathyroid	0	0	0	0	0	0	0	0	35	0
Peritoneum	0	0	0	0	0	0	0	0	11	0
Pituitary	0	0	0	0	0	0	0	0	57	0
Placenta	0	0	0	0	0	0	0	0	2	0
Pleura	0	0	0	0	0	0	0	0	264	0
Prostate	0	0	3	3	0	37	10	0	2078	0.289
Salivary gland	0	0	0	0	0	0	0	0	80	0
Skin	0	0	13	31	0	22	6	0	1305	3.372
Small intestine	0	0	0	1	1	0	0	0	53	3.774
Soft tissue	0	0	0	0	0	5	0	0	567	0
Stomach	0	0	8	10	0	40	0	0	872	2.064
Testis	0	0	0	0	0	0	0	0	169	0
Thymus	0	0	0	0	0	0	0	0	152	0
Thyroid	0	0	0	9	2	0	0	0	942	1.168
Upper aerodigestive tract	0	0	1	5	0	4	0	0	1296	0.463
Urinary tract	0	0	2	9	0	10	0	0	679	1.62
Vulva	0	0	0	0	0	0	0	0	3	0

##NS: unknown type

**Table S3.** TCGA and GTEx samples.

Cancer_type	Normal.mean	Tumor.mean	Logfc	Pvalue	Normal num	Tumor num	Detail
ACC	-1.308843	0.8991506	2.2079936	8.23E-11	128	77	Adrenocortical carcinoma
BLCA	-0.3509214	4.9707316	5.321653	6.59E-15	28	407	Bladder urothelial carcinoma
BRCA	-0.6600106	4.0017489	4.6617595	7.33E-137	292	1098	Breast invasive carcinoma
CESC	-3.7105923	5.8869425	9.5975348	1.03E-09	13	306	Cervical squamous cell carcinoma and endocervical adenocarcinoma
CHOL	-2.9221333	3.1901806	6.1123139	4.59E-06	9	36	Cholangio carcinoma
COAD	-0.8951401	5.2980245	6.1931646	2.45E-99	349	288	Colon adenocarcinoma
DLBC	6.1310729	6.8949915	0.7639186	7.81E-08	107	47	Lymphoid neoplasm diffuse large B-cell lymphoma
ESCA	-1.0560677	5.6398956	6.6959633	3.01E-92	668	182	Esophageal carcinoma
GBM	-5.0849353	4.0848572	9.1697924	9.03E-62	207	165	Glioblastoma multiforme
HNSC	2.4676227	5.2156567	2.748034	5.89E-26	44	520	Head and neck squamous cell carcinoma
KICH	-2.6252189	-1.0309288	1.5942901	3.40E-05	53	66	Kidney chromophobe
KIRC	-2.766057	0.9112407	3.6772977	5.07E-39	100	531	Kidney renal clear cell carcinoma
KIRP	-2.7614833	0.6606356	3.422119	3.95E-21	60	289	Kidney renal papillary cell carcinoma
LAML	0.6936217	4.335174	3.6415523	1.16E-65	337	173	Acute myeloid leukemia
LGG	-5.0849353	0.8013128	5.886248	3.50E-89	207	522	Brain lower grade glioma
LIHC	-2.3198356	2.0033752	4.3232109	3.01E-56	160	371	Liver hepatocellular carcinoma
LUAD	0.2917239	3.8211272	3.5294032	1.23E-106	347	515	Lung adenocarcinoma
LUSC	0.3090672	5.3213818	5.0123146	2.16E-130	338	498	Lung squamous cell carcinoma
MESO	NA	3.4563218	NA	NA	NA	87	Mesothelioma
OV	-3.3205159	4.8593319	8.1798478	8.08E-49	88	427	Ovarian serous cystadenocarcinoma
PAAD	1.7998339	2.9644302	1.1645962	9.39E-27	171	179	Pancreatic adenocarcinoma
PCPG	-2.1921667	-1.2632659	0.9289007	1.83E-01	3	182	Pheochromocytoma and Paraganglioma
PRAD	-1.3854638	1.0329685	2.4184323	4.93E-50	152	496	Prostate adenocarcinoma
READ	-1.3023349	5.4936087	6.7959436	5.15E-47	318	92	Rectum adenocarcinoma
SARC	-0.64105	3.8995057	4.5405557	1.80E-02	2	262	Sarcoma
SKCM	0.2628796	4.3137797	4.0509002	1.74E-158	558	469	Skin cutaneous melanoma
STAD	-0.408955	5.0930459	5.5020009	3.07E-85	211	414	Stomach adenocarcinoma

Cancer_type	Normal.mean	Tumor.mean	Logfc	Pvalue	Normal num	Tumor num	Detail
TGCT	4.2169406	8.0783182	3.8613776	7.45E-46	165	137	Testicular germ cell tumors
THCA	-1.6424183	-0.0863418	1.5560765	8.46E-55	338	512	Thyroid carcinoma
THYM	5.09365	4.4921403	-0.6015097	6.47E-01	2	119	Thymoma
UCEC	-3.1823418	5.0303983	8.2127401	4.17E-41	91	181	Uterine corpus endometrial carcinoma
UCS	-3.5328526	5.4762439	9.0090964	3.93E-23	78	57	Uterine carcinosarcoma
UVM	NA	0.8823873	NA	NA	NA	79	Uveal melanoma

**Table S4.** Prognoscan database meta-analysis.

Caner type	Sbotype	N	Cox P-value	HR	Endpoint	Dataset	Probe ID
Bladder cancer		30	1.89E-02	4.884222	Overall survival	GSE5287	201710_at
Bladder cancer		30	4.27E-02	0.3369907	Overall survival	GSE5287	216421_at
Blood cancer	AML	163	4.02E-02	1.9446811	Overall survival	GSE12417-GPL96	216421_at
Blood cancer	Multiple myeloma	559	5.34E-04	1.350538	Disease specific survival	GSE2658	201710_at
Brain cancer	Astrocytoma	77	2.45E-04	1.6398553	Overall survival	GSE4271-GPL96	201710_at
Brain cancer	Glioma	50	1.46E-02	1.392811	Overall survival	MGH-glioma	1854_at
Brain cancer	Glioma	74	2.76E-03	1.5781369	Overall survival	GSE4412-GPL96	201710_at
Breast cancer		158	1.30E-02	1.7626758	Overall survival	GSE3143	1854_at
Breast cancer		204	3.84E-03	1.2628221	Relapse free survival	GSE12276	201710_at
Breast cancer		77	1.66E-03	2.6244701	Distant metastasis free survival	GSE9195	201710_at
Breast cancer		77	4.07E-03	2.2250936	Relapse free survival	GSE9195	201710_at

Caner type	Sbutype	N	Cox P-value	HR	Endpoint	Dataset	Probe ID
Breast cancer		136	1.46E-02	1.7433299	Distant metastasis free survival	GSE12093	201710_at
Breast cancer		200	2.61E-02	1.4009086	Distant metastasis free survival	GSE11121	201710_at
Breast cancer		60	3.91E-02	2.1735853	Relapse free survival	GSE1379	3776
Breast cancer		286	9.07E-03	1.3422899	Distant metastasis free survival	GSE2034	201710_at
Breast cancer		159	2.52E-03	1.7864082	Relapse free survival	GSE1456-GPL96	201710_at
Breast cancer		159	2.46E-03	2.0273165	Disease specific survival	GSE1456-GPL96	201710_at
Breast cancer		159	7.42E-03	1.6504764	Overall survival	GSE1456-GPL96	201710_at
Breast cancer		236	3.02E-04	1.6169619	Disease specific survival	GSE3494-GPL96	201710_at
Breast cancer		249	1.25E-05	1.5616719	Disease free survival	GSE4922-GPL96	201710_at
Breast cancer		198	3.07E-02	1.1868895	Distant metastasis free survival	GSE7390	201710_at
Breast cancer		198	1.18E-02	1.235614	Overall survival	GSE7390	201710_at
Breast cancer		198	2.02E-02	1.1599443	Relapse free survival	GSE7390	201710_at
Lung cancer	Adenocarcinoma	82	4.42E-02	2.4170239	Overall survival	jacob-00182-CANDF	216421_at

Caner type	Sbutype	N	Cox P-value	HR	Endpoint	Dataset	Probe ID
Lung cancer	Adenocarcinoma	82	2.38E-02	2.6443334	Overall survival	jacob-00182-CANDF	201710_at
Lung cancer	Adenocarcinoma	104	1.32E-02	1.9670402	Overall survival	jacob-00182-MSK	201710_at
Lung cancer	Adenocarcinoma	117	7.21E-03	2.3988657	Overall survival	GSE13213	A_23_P143190
Lung cancer	Adenocarcinoma	204	3.24E-04	1.6675492	Overall Survival	GSE31210	201710_at
Lung cancer	Adenocarcinoma	204	4.74E-07	1.6769791	Relapse free survival	GSE31210	201710_at
Lung cancer	NSCLC	50	8.09E-03	3.8364981	Overall survival	GSE4716-GPL3694	546
Ovarian cancer		278	3.56E-02	1.2934059	Overall survival	GSE9891	201710_at
Ovarian cancer		133	2.80E-03	2.4199479	Overall survival	DUKE-OC	216421_at
Ovarian cancer		81	2.08E-05	4.5475562	Overall survival	GSE8841	8776
Prostate cancer		281	9.09E-04	1.4044531	Overall survival	GSE16560	DAP1_5244
Skin cancer	Melanoma	38	1.05E-03	13.2496349	Overall survival	GSE19234	201710_at
Soft tissue cancer	Liposarcoma	140	3.67E-02	0.3882763	Distant recurrence free survival	GSE30929	216421_at
Soft tissue cancer	Liposarcoma	140	1.06E-04	5.981432	Distant recurrence free survival	GSE30929	201710_at

**Table S5.** KM plotter about 5-years overall survival.

Tumor	survp	HR(High/Low)	95% CI	Coxpvalue
Breast	p < 0.0001	2.24	1.7 to 2.95	0
Gastric	p = 0.03155	1.21	1.02 to 1.44	0.032
Lung	p < 0.0001	2.1	1.82 to 2.42	0
Ovary	p = 0.04025	1.16	1.01 to 1.33	0.04



**Table S6.** TCGA 5-years overall survival.

Cancer_type	survp	HR (high/low)	95% CI	Coxpvalue
ACC	p < 0.0001	6.45	2.38 to 17.46	0
KIRC	p < 0.0001	2.01	1.43 to 2.82	0
KIRP	p < 0.0001	4.33	1.99 to 9.45	0
LGG	p < 0.0001	3.68	2.31 to 5.85	0
MESO	p < 0.0001	4	2.4 to 6.64	0
LIHC	p = 0.00152	1.79	1.24 to 2.59	0.002
BRCA	p = 0.00313	1.84	1.22 to 2.77	0.004
SKCM	p = 0.00362	1.63	1.17 to 2.28	0.004
SARC	p = 0.00799	1.78	1.16 to 2.75	0.009
PAAD	p = 0.01386	1.68	1.11 to 2.56	0.015
THYM	p = 0.02255	0	0 to Inf	0.999
KICH	p = 0.03379	7.01	0.86 to 57	0.069
LAML	p = 0.05794	1.45	0.98 to 2.13	0.061
UVM	p = 0.05954	2.3	0.94 to 5.59	0.067
LUAD	p = 0.11687	1.27	0.94 to 1.72	0.118
PCPG	p = 0.20999	2.73	0.53 to 14.11	0.229
DLBC	p = 0.2161	0.37	0.07 to 1.91	0.235
CESC	p = 0.2396	0.74	0.45 to 1.22	0.241
OV	p = 0.37739	1.13	0.86 to 1.47	0.377
UCEC	p = 0.39069	1.35	0.67 to 2.73	0.394
COAD	p = 0.41777	1.24	0.74 to 2.06	0.419
HNSC	p = 0.46174	1.11	0.84 to 1.46	0.462
ESCA	p = 0.49509	1.17	0.75 to 1.84	0.495
PRAD	p = 0.50216	1.77	0.32 to 9.7	0.508
STAD	p = 0.52553	0.9	0.66 to 1.24	0.525
TGCT	p = 0.56006	2.01	0.18 to 22.2	0.568
THCA	p = 0.6729	1.26	0.44 to 3.62	0.674
GBM	p = 0.83721	1.04	0.73 to 1.47	0.823
READ	p = 0.87601	0.92	0.35 to 2.4	0.864
CHOL	p = 0.92245	0.95	0.36 to 2.49	0.922
LUSC	p = 0.92676	1.01	0.76 to 1.36	0.927
UCS	p = 0.95757	0.98	0.49 to 1.97	0.96
BLCA	p = 0.97197	1.01	0.74 to 1.36	0.972

**Table S7.** TCGA 5-years Disease-specific survival.

Cancer_type	Survival	HR (high/low)	95% CI	Coxpvalue
ACC	p < 0.0001	7.29	2.45 to 21.74	0
KIRC	p < 0.0001	2.58	1.66 to 4	0
KIRP	p < 0.0001	13.88	3.29 to 58.62	0
LGG	p < 0.0001	4.14	2.5 to 6.86	0
MESO	p < 0.0001	4.45	2.33 to 8.49	0
BRCA	p = 0.00039	2.66	1.51 to 4.66	0.001
LIHC	p = 0.00043	2.33	1.43 to 3.79	0.001
SKCM	p = 0.00747	1.62	1.13 to 2.3	0.008
KICH	p = 0.01479	Inf	0 to Inf	0.998
PAAD	p = 0.01847	1.75	1.09 to 2.8	0.02
LUAD	p = 0.02559	1.55	1.05 to 2.28	0.027
UVM	p = 0.03454	2.62	1.04 to 6.64	0.042
SARC	p = 0.03504	1.65	1.03 to 2.63	0.037
PRAD	p = 0.0973	Inf	0 to Inf	0.999
READ	p = 0.13829	4.54	0.51 to 40.77	0.177
PCPG	p = 0.14571	4.42	0.49 to 39.56	0.184
THYM	p = 0.16276	0	0 to Inf	0.999
CESC	p = 0.22609	0.71	0.41 to 1.24	0.228
ESCA	p = 0.2966	1.34	0.77 to 2.33	0.298
LUSC	p = 0.29843	1.27	0.81 to 1.99	0.3
DLBC	p = 0.31097	0.33	0.03 to 3.17	0.335
OV	p = 0.52223	1.1	0.82 to 1.46	0.522
TGCT	p = 0.56006	2.01	0.18 to 22.2	0.568
BLCA	p = 0.62104	1.1	0.76 to 1.58	0.621
UCS	p = 0.65309	1.18	0.57 to 2.45	0.651
STAD	p = 0.66259	1.1	0.73 to 1.65	0.663
THCA	p = 0.7628	1.26	0.28 to 5.63	0.763
HNSC	p = 0.82802	1.04	0.73 to 1.47	0.829
COAD	p = 0.85801	0.94	0.47 to 1.88	0.858
CHOL	p = 0.93069	0.96	0.34 to 2.66	0.931
GBM	p = 0.97723	1	0.69 to 1.45	0.995
UCEC	p = 0.99041	1	0.44 to 2.28	0.993

**Table S8.** TCGA 5-years progressive-free interval.

Cancer_type	survp	HR (high/low)	95% CI	Coxpvalue
ACC	p < 0.0001	4.08	2 to 8.33	0
KIRC	p < 0.0001	2.22	1.57 to 3.14	0
KIRP	p < 0.0001	3.33	1.8 to 6.15	0
PRAD	p = 0.00026	2.26	1.44 to 3.53	0
MESO	p = 0.00032	2.67	1.54 to 4.63	0
LGG	p = 0.00042	1.69	1.26 to 2.26	0
LIHC	p = 0.00086	1.64	1.22 to 2.21	0.001
UVM	p = 0.00153	3.81	1.57 to 9.24	0.003
KICH	p = 0.00247	Inf	0 to Inf	0.998
THCA	p = 0.0034	2.3	1.3 to 4.09	0.004
PAAD	p = 0.01265	1.63	1.11 to 2.41	0.013
BRCA	p = 0.01744	1.55	1.08 to 2.23	0.018
SARC	p = 0.01933	1.5	1.07 to 2.11	0.02
PCPG	p = 0.02265	2.91	1.11 to 7.59	0.029
TGCT	p = 0.02491	2.31	1.09 to 4.91	0.029
SKCM	p = 0.06275	1.27	0.99 to 1.63	0.063
UCEC	p = 0.08624	1.68	0.92 to 3.07	0.09
LUAD	p = 0.17024	1.21	0.92 to 1.6	0.17
READ	p = 0.23213	1.75	0.69 to 4.46	0.238
UCS	p = 0.28186	1.44	0.74 to 2.8	0.287
GBM	p = 0.31088	0.84	0.59 to 1.18	0.314
CHOL	p = 0.38642	1.48	0.6 to 3.64	0.389
LUSC	p = 0.3989	1.16	0.82 to 1.63	0.4
ESCA	p = 0.51412	1.15	0.75 to 1.76	0.512
THYM	p = 0.62246	1.26	0.5 to 3.2	0.623
CESC	p = 0.83413	0.95	0.6 to 1.52	0.835
STAD	p = 0.85148	1.03	0.74 to 1.45	0.851
COAD	p = 0.87462	1.04	0.67 to 1.61	0.874
HNSC	p = 0.8784	1.02	0.77 to 1.36	0.879
DLBC	p = 0.90395	1.08	0.29 to 4.06	0.904
OV	p = 0.90597	1.01	0.81 to 1.27	0.91
BLCA	p = 0.93614	1.01	0.75 to 1.37	0.935

**Table S9.** Permutation test.

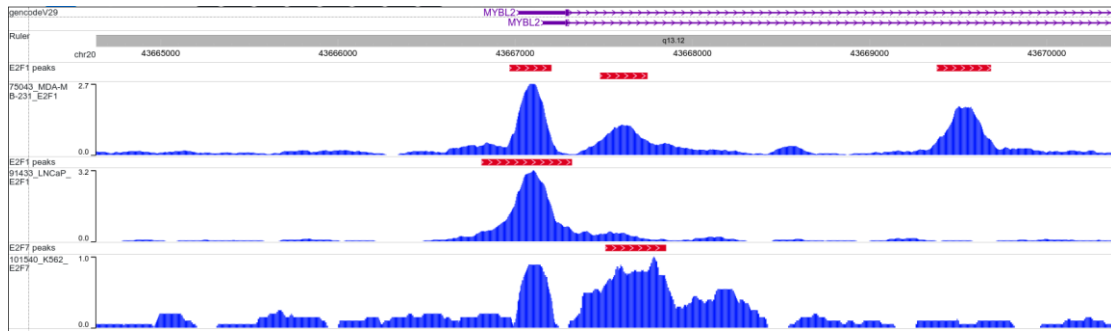
PRAD permutation test	BRCA permutation test
Approximative wilcoxon-pratt signed-rank test data: y by x (pos, neg) stratified by block Z = - 4.7776, p-value < 2.2E-16 alternative hypothesis: true mu is not equal to 0	Approximative wilcoxon-pratt signed-rank test data: y by x (pos, neg) stratified by block Z = 26.03, p-value < 2.2E-16 alternative hypothesis: true mu is not equal to 0

**LUAD Permutation test.**

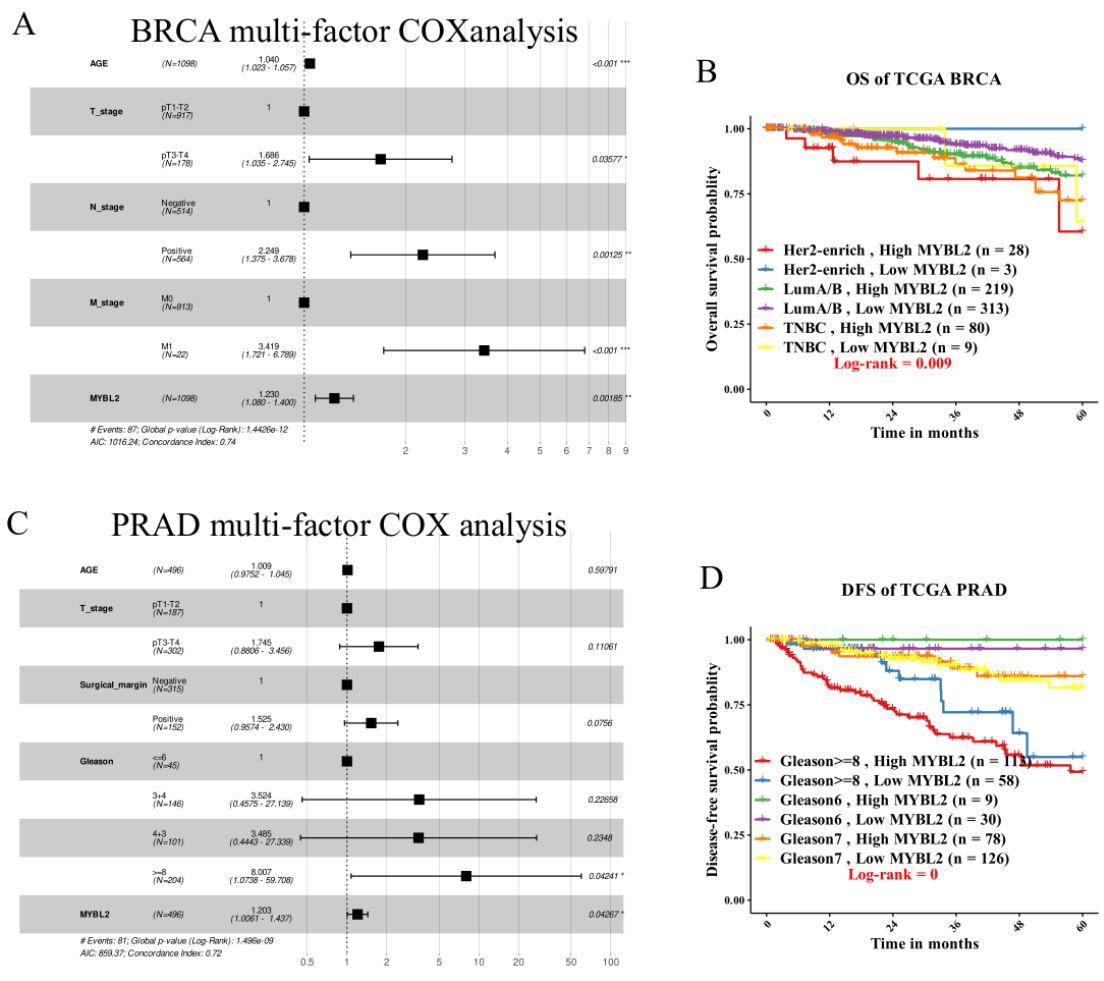
Approximative wilcoxon-pratt signed-rank test  
Data: y by x (pos, neg)  
Stratified by block  
Z = 11.862, p-value < 2.2E-16  
Alternative hypothesis: true mu is not equal to 0

**Table S10.** Potential TF binding to MYBL2 promotor.

	Potential MYBL2 promoter binding sequence	Start	End	Width	Score
E2F7	[ATGATCGCGAAAC]	1002	1014	13	0.86
E2F2	[ATCGCGAAAC]	1005	1014	10	0.87
E2F1	[GCCTGGCGACAAAG]	859	872	14	0.83
E2F8	NONE	NONE	NONE	NONE	NONE
FOXM1	NONE	NONE	NONE	NONE	NONE
ZNF695	[AGGATTA]	47	53	7	0.85
	[AAGGTAA]	101	107	7	0.97
	[ATGGTTT]	441	447	7	0.85
	[ATGGTGA]	683	689	7	0.84
	[AGGGTCT]	1091	1097	7	0.98
	[AGGGTGA]	1135	1141	7	0.86
	[AGGGTCT]	1389	1395	7	0.98



**Figure S1.** The actual motif occupancy of E2F1 and E2F7 on the promoter region of MYBL2 using the Chip-seq data in cancer cells from the Cristrome database.



**Figure S2.** (A,C) Multi-factor analysis in PRAD and BRCA. (B,D) Survival analysis combined molecular subtypes and MYBL2 on BRCA and PRAD.