

Supplemental Table 1. Antibodies for immunohistochemical staining

Antigen	Antibody Clone	Manufacturer
CD20	L26	Leica
BCL2	bcl2/100/D5	Novocastra
MYC	EP121	Genetex
CD10	56C6	Leica
BCL6	LN22	Novocastra
MUM1	MUM1p	DAKO
CD3	LN10	Leica
CD4	IF6	Nichirei
CD8	C8/144B	Nichirei
CD31	1A10	Leica
CD68	PGM1	DAKO

Supplemental Table 2. FISH probes

Probe Name	Manufacturer	Catalogue Number
Vysis LSI MYC Dual Color	Abbott	5J91-01
Vysis LSI BCL2 Dual Color	Abbott	7J75-01

Abbreviations: FISH: fluorescent in situ hybridization

Supplemental Table 4. Patient characteristics of final cohort and excluded cohort

	Final cohort	Excluded cohort	<i>P</i> value
Number of patients	1050	526	
Age (n, %)			
Median	72	72	
<61	183 (17.4)	104 (19.8)	0.268
≥61	867 (82.6)	422 (80.2)	
Sex (n, %)			
Male	591 (56.3)	290 (55.1)	0.667
Female	459 (43.7)	236 (44.9)	
Performance Status (n, %)			
0-1	853 (82.4)	382 (79.6)	0.2
2-4	182 (17.6)	98 (20.4)	
NA	15	46	
Stage (n, %)			
I/II	472 (45)	223 (45.1)	1
III/IV	578 (55)	272 (54.9)	
NA	0	31	
LDH (n, %)			
Normal	399 (38)	179 (37.4)	0.865
>ULN	650 (62)	299 (62.6)	
NA	1	48	
Extranodal sites (n, %)			
0-1	830 (79)	377 (77.9)	0.639
≥2	220 (21)	107 (22.1)	
NA	0	42	
Biopsy site (n, %)			
Nodal	560 (53.3)	179 (34.4)	<0.0001
Extranodal	490 (46.7)	342 (65.6)	
IPI risk group (n, %)			
Low (0-1)	289 (27.9)	143 (30.6)	0.241
Low-Intermediate (2)	264 (25.5)	106 (22.7)	
High-Intermediate (3)	253 (24.5)	101 (21.6)	
High (4-5)	228 (22.1)	117 (25.1)	
NA	16	59	

Abbreviations: LDH: lactate dehydrogenase, IPI: International Prognostic Index,
NA: not available, ULN: upper limit of normal

Supplemental Table 5. Patient characteristics of the present cohort and BCC cohort

	Present cohort	BCC	P value
Number of patients	1050	804	
Age (n, %)			
Median	72	66	
<61	183 (17.4)	239 (29.7)	<0.001
≥61	867 (82.6)	565 (70.3)	
Sex (n, %)			
Male	591 (56.3)	453 (56.3)	1
Female	459 (43.7)	351 (43.7)	
Performance Status (n, %)			
0-1	853 (82.4)	447 (61.5)	<0.001
2-4	182 (17.6)	280 (38.5)	
NA	15	77	
Stage (n, %)			
I/II	472 (45)	345 (46.2)	0.631
III/IV	578 (55)	402 (53.8)	
NA	0	57	
LDH (n, %)			
Normal	399 (38)	348 (53.3)	<0.001
>ULN	650 (62)	305 (46.7)	
NA	1	151	
Extranodal sites (n, %)			
0-1	830 (79)	591 (79.1)	1
≥2	220 (21)	156 (20.9)	
NA	0	57	
IPI risk group (n, %)			
Low (0-1)	289 (27.9)	225 (34.1)	0.104
Low-Intermediate (2)	264 (25.5)	157 (23.8)	
High-Intermediate (3)	253 (24.5)	141 (21.4)	
High (4-5)	228 (22.1)	136 (20.6)	
NA	16	145	

Abbreviations: BCC: BC Cancer, LDH: lactate dehydrogenase,

IPI: International Prognostic Index, NA: not available, ULN: upper limit of normal

Supplemental Table 6. Difference of biopsies obtained from nodal or extranodal sites according to refined cell-of-origin

	GCB	DZsig ^{pos}	ABC	UNC	P value
Biopsy site (n, %)					
Nodal	170 (46.3)	39 (60)	270 (57.7)	81 (54)	0.0074
Extranodal	197 (53.7)	26 (40)	198 (42.3)	69 (46)	

Abbreviations: ABC: activated B-cell-like, DZsig^{pos}: dark zone signature positive, GCB: germinal center B-cell-like, UNC: unclassified

Supplemental Table 7. Univariable analysis including refined cell-of-origin and IPI groups

Variables	Progression free survival		Overall survival	
	HR (95% CI)	P value	HR (95% CI)	P value
DZsig ^{pos} vs GCB	2.96 (2.01-4.36)	<0.001	2.80 (1.74-4.50)	<0.001
ABC vs GCB	1.99 (1.56-2.54)	<0.001	2.12 (1.57-2.86)	<0.001
UNC vs GCB	1.69 (1.23-2.31)	0.001	1.66 (1.12-2.46)	0.012
IPI (High/High-intermediate vs Low/Low-intermediate)	2.75 (2.24-3.39)	<0.001	2.587 (2.01-3.33)	<0.001

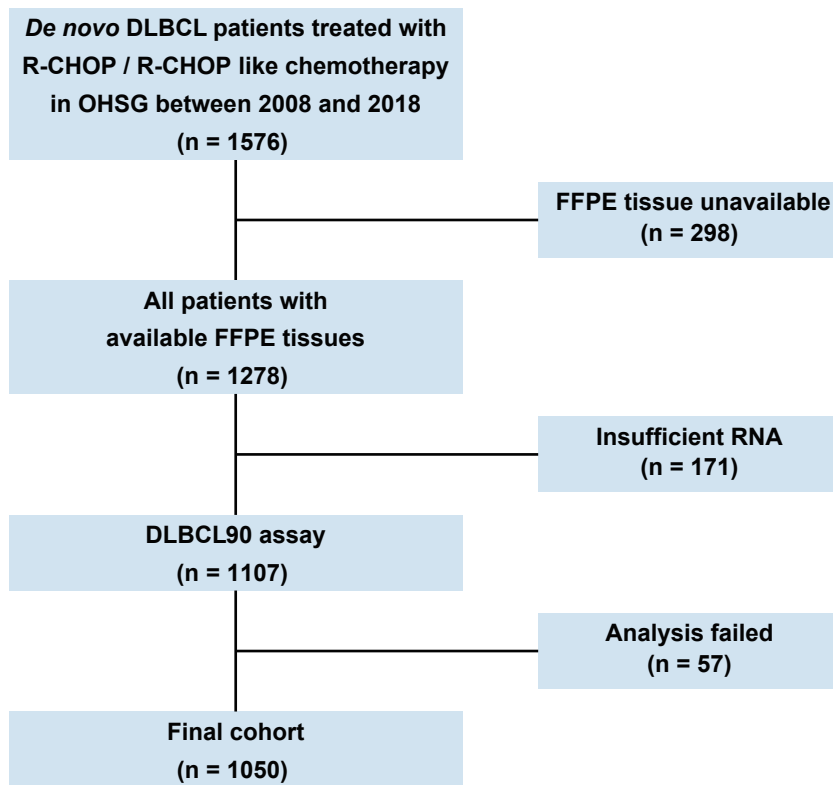
Abbreviations: ABC: activated B-cell-like, DZsig^{pos}: dark zone signature positive, GCB: germinal center B-cell-like, UNC: unclassified, IPI: International Prognostic Index

Supplemental Table 8. Primary refractory status according to refined cell-of-origin

	GCB	DZsig ^{pos}	ABC	UNC
Number of patients	72	28	164	52
Relapse timing (n, %)				
<9 months	32 (44.4)	19 (67.9)	81 (49.4)	21 (40.4)
≥9 months	40 (55.6)	9 (32.1)	83 (50.6)	31 (59.6)

Abbreviations: ABC: activated B-cell-like, DZsig^{pos}: dark zone signature positive, GCB: germinal center B-cell-like, UNC: unclassified

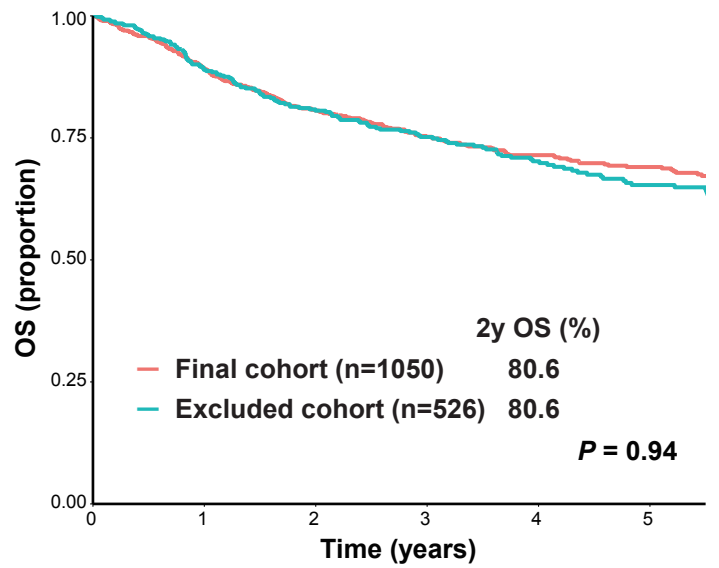
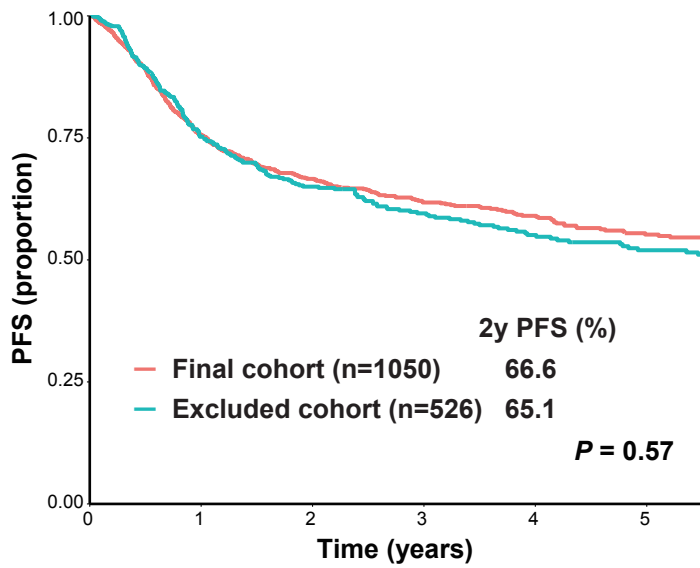
Supplemental Figure 1



Supplemental Figure 1: Flowchart of the diffuse large B-cell lymphoma study cohort.

FFPE: formalin-fixed paraffin-embedded, OHSG: Okayama hematology study group, R-CHOP: rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone

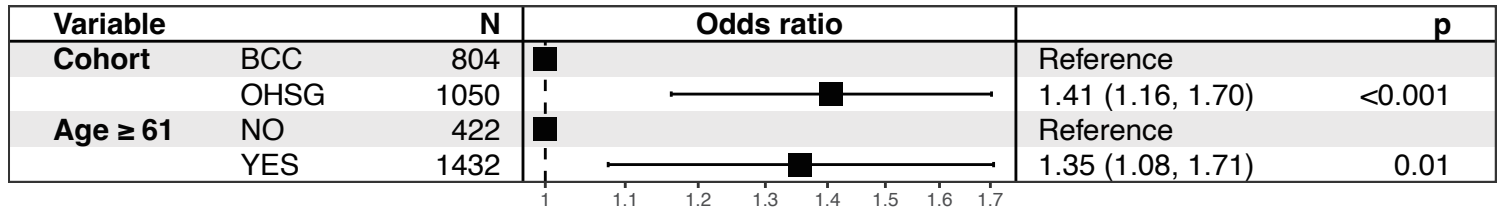
Supplemental Figure 2



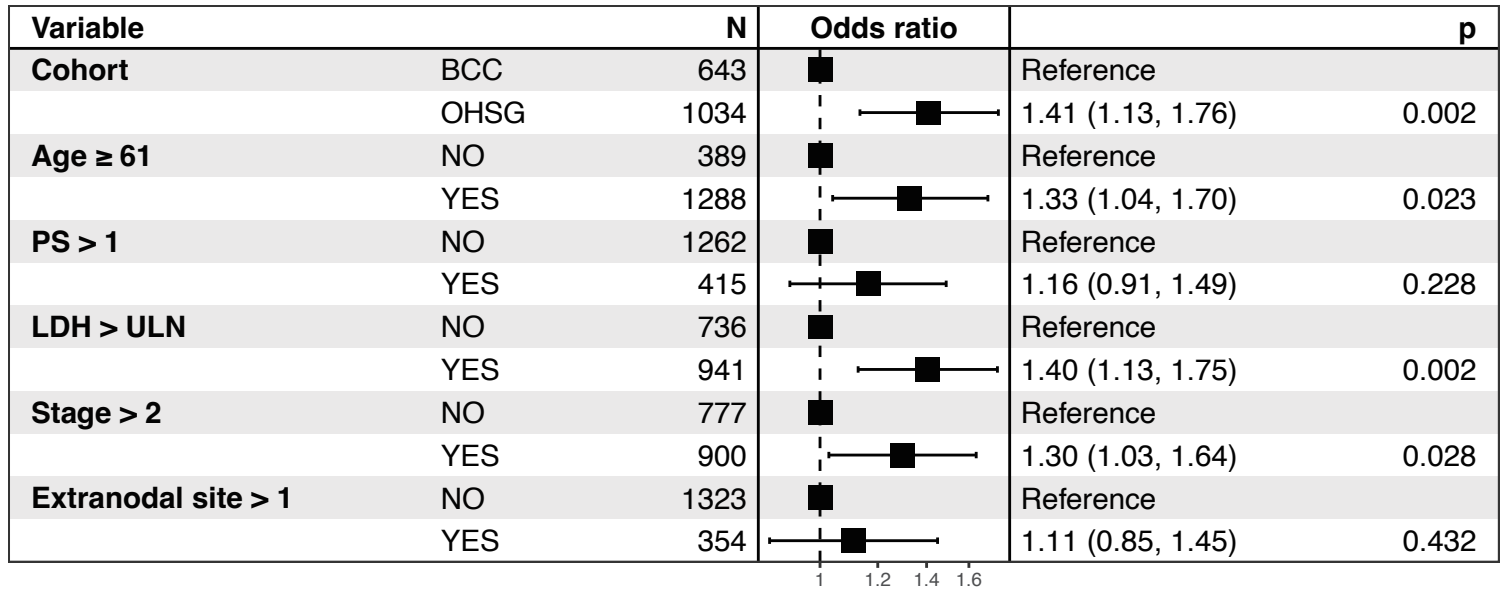
Supplemental Figure 2: Clinical outcomes of patients with diffuse large B-cell lymphoma stratified by the availability of gene expression profiling (GEP) data. Final cohort includes patients with GEP data, and excluded cohort consists of patients who were not included due to reasons shown in Figure S1.

Supplemental Figure 3

Model 1

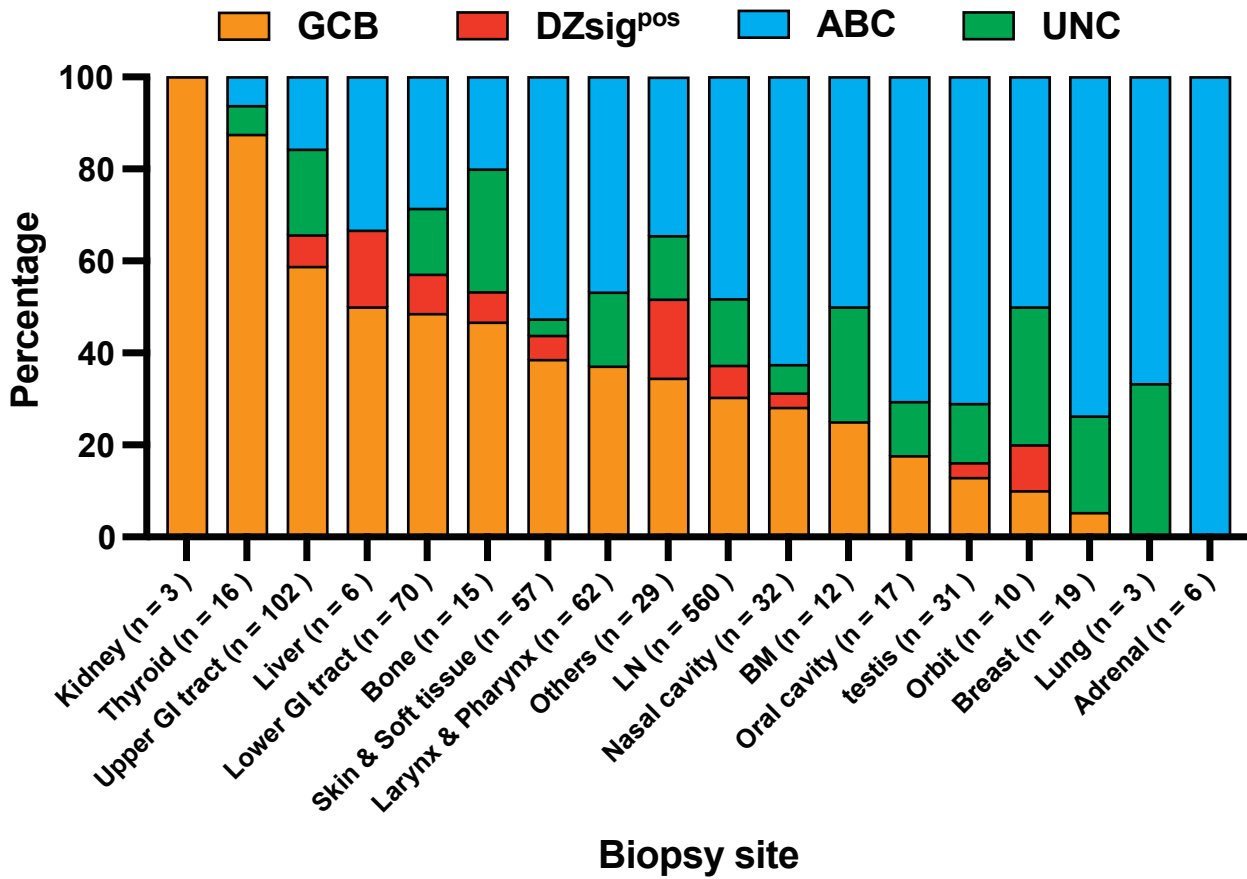


Model 2



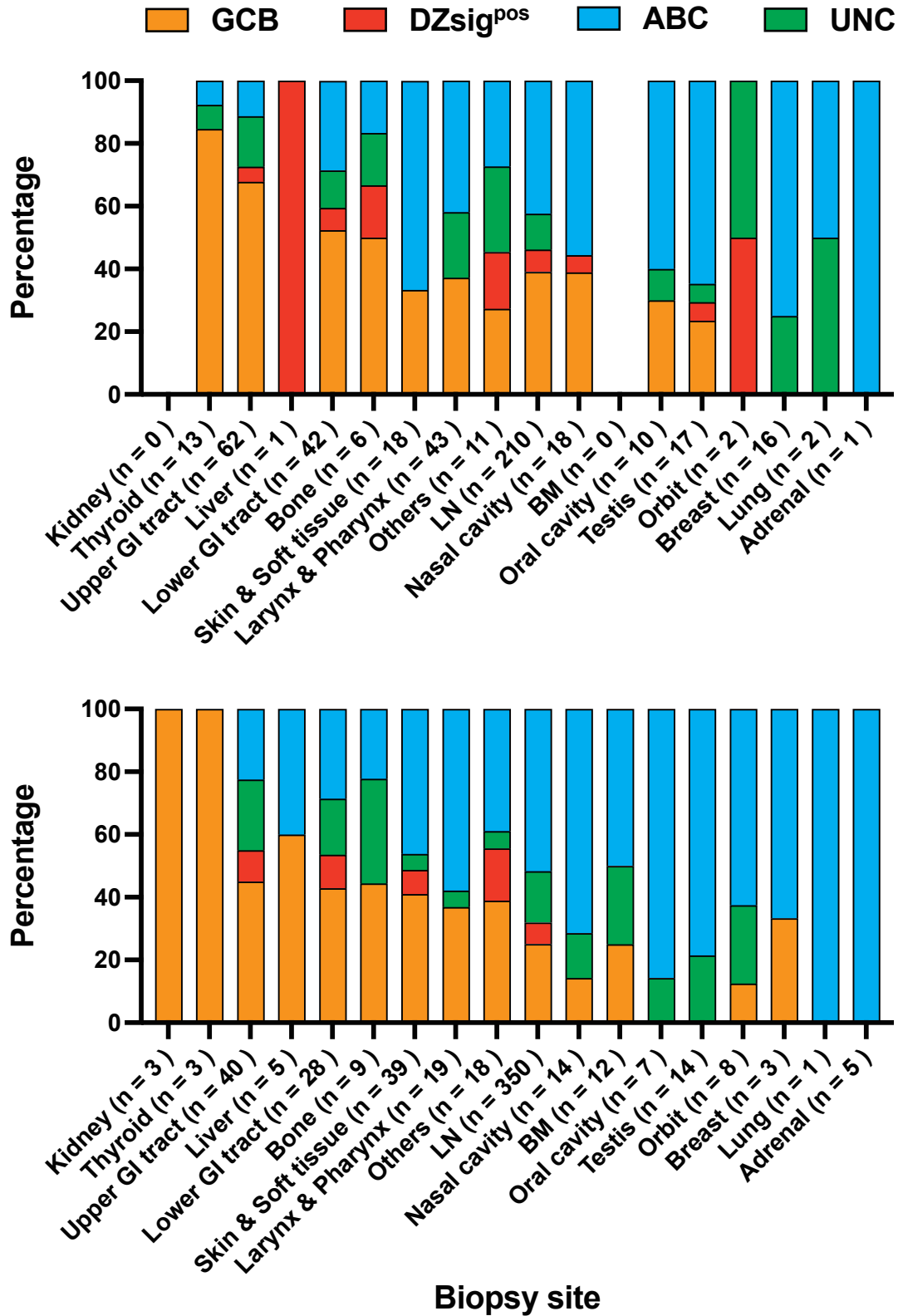
Supplemental Figure 3 : Associations of the COO distributions in OHSG and BCC cohorts with clinical factors based on multivariable logistic regression models. Forest plots showing the comparison of odds ratios adjusted for the cohorts and age group (top), and for the cohort, age group and other IPI factors (bottom). COO: cell-of-origin, BCC: BC Cancer, OHSG: Okayama Hematology Study Group

Supplemental Figure 4



Supplemental Figure 4: Distribution of refined cell-of-origin (COO) according to biopsy sites. Bar plots represent the frequency of refined COO in each biopsy site. ABC: activated B-cell-like, DZsig: dark zone signature, GCB: germinal center B-cell-like, pos: positive, UNC: unclassified, GI tract: gastrointestinal tract

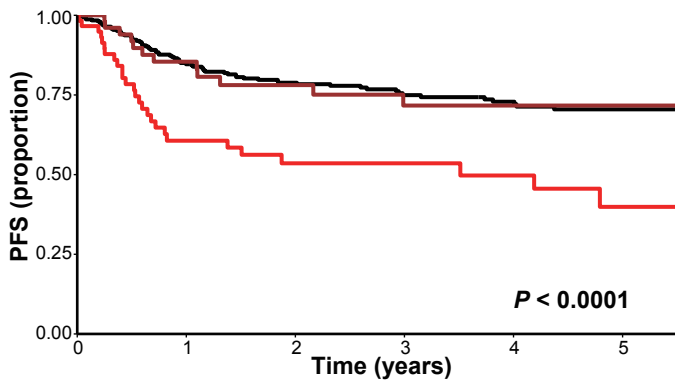
Supplemental Figure 5



Supplemental Figure 5: Distribution of refined cell-of-origin (COO) according to biopsy sites. Bar plots represent the frequency of refined COO of patients with stage I/II (top) and III/IV (bottom) in each biopsy site. ABC: activated B-cell-like, DZsig: dark zone signature, GCB: germinal center B-cell-like, pos: positive, UNC: unclassified, GI tract: gastrointestinal tract

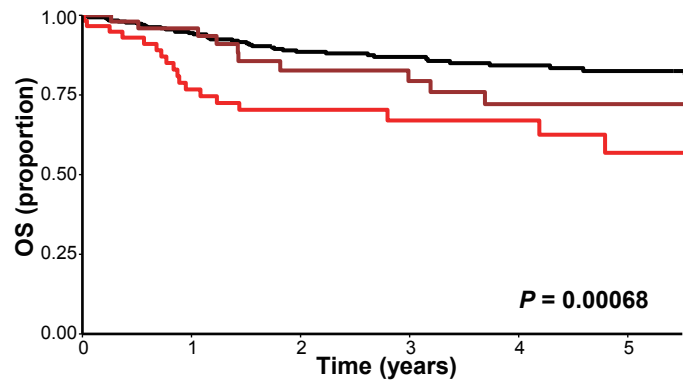
Supplemental Figure 6

A



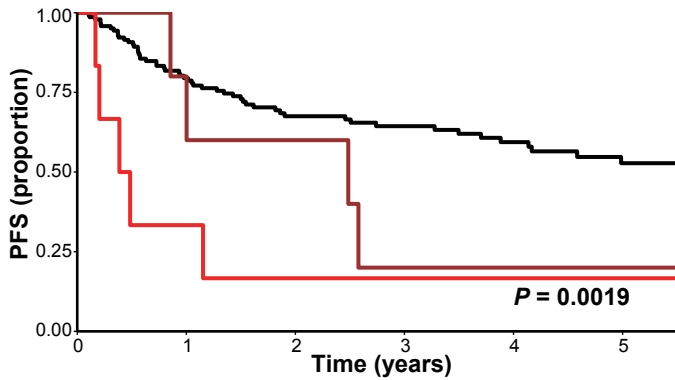
Number at risk							
—	313	221	172	122	95	73	
—	54	36	26	21	18	15	
—	59	29	20	15	12	6	
—	DZsig ^{neg} (n = 313)		2y-PFS (95% CI)		HR (95% CI)		
—	DZsig ^{ind} (n = 54)		78.8% (74.1-83.9)		1		
—	DZsig ^{pos} (n = 59)		78.1% (66.9-91.2)		0.99 (0.56-1.75)		
			53.5% (41.3-69.5)		2.71 (1.78-4.13)		

B



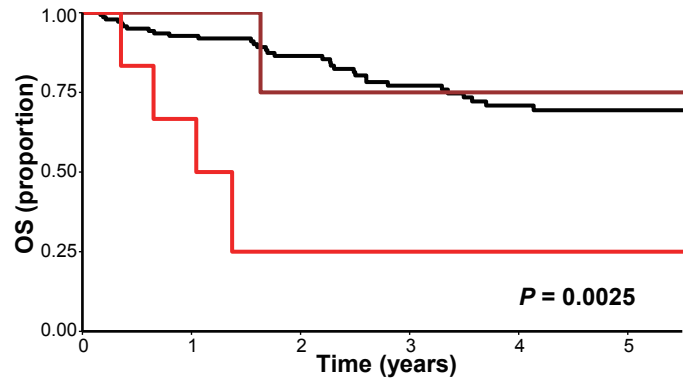
Number at risk							
—	313	245	192	143	110	86	
—	54	41	28	24	19	15	
—	59	37	26	19	15	9	
—	DZsig ^{neg} (n = 313)		2y-OS (95% CI)		HR (95% CI)		
—	DZsig ^{ind} (n = 54)		88.5% (84.7-92.5)		1		
—	DZsig ^{pos} (n = 59)		82.7% (71.6-95.4)		1.47 (0.78-2.76)		
			70.3% (58.7-84.3)		2.67 (1.58-4.50)		

C



Number at risk							
—	145	103	71	57	42	27	
—	5	4	3	1	1	1	
—	6	2	1	1	1	1	
—	DZsig ^{neg} (n = 145)		2y-PFS (95% CI)		HR (95% CI)		
—	DZsig ^{ind} (n = 5)		67.5% (59.8-76.2)		1		
—	DZsig ^{pos} (n = 6)		60.0% (29.3-100.0)		2.05 (0.74-5.67)		
			16.7% (2.8-99.7)		4.25 (1.69-10.7)		

D

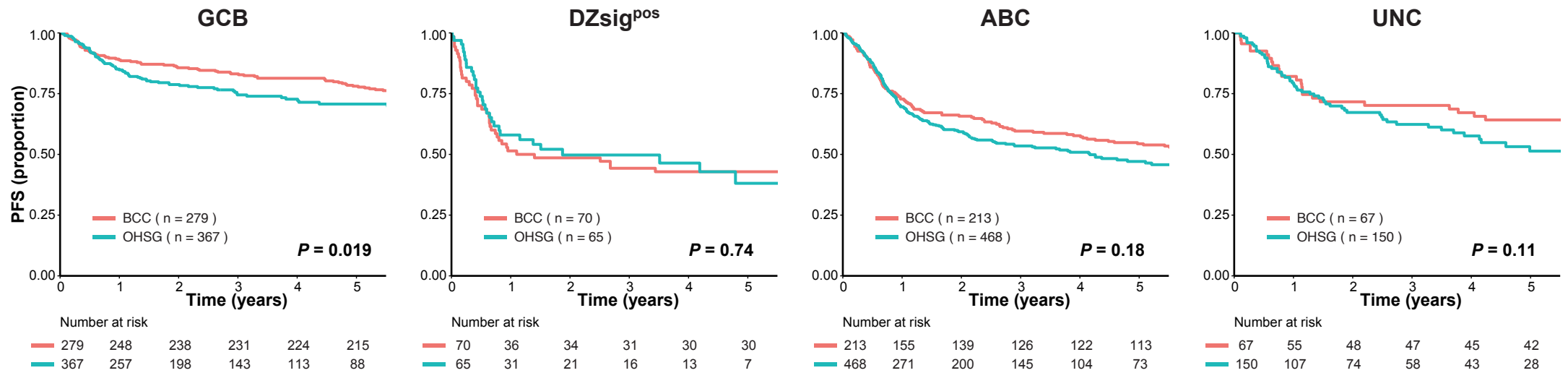


Number at risk							
—	145	119	89	66	50	36	
—	5	5	3	3	2	1	
—	6	4	1	1	1	1	
—	DZsig ^{neg} (n = 145)		2y-OS (95% CI)		HR (95% CI)		
—	DZsig ^{ind} (n = 5)		86.4% (80.6-92.7)		1		
—	DZsig ^{pos} (n = 6)		75.0% (42.6-100.0)		0.76 (0.10-5.59)		
			25.0% (5.0-100.0)		5.18 (1.81-14.8)		

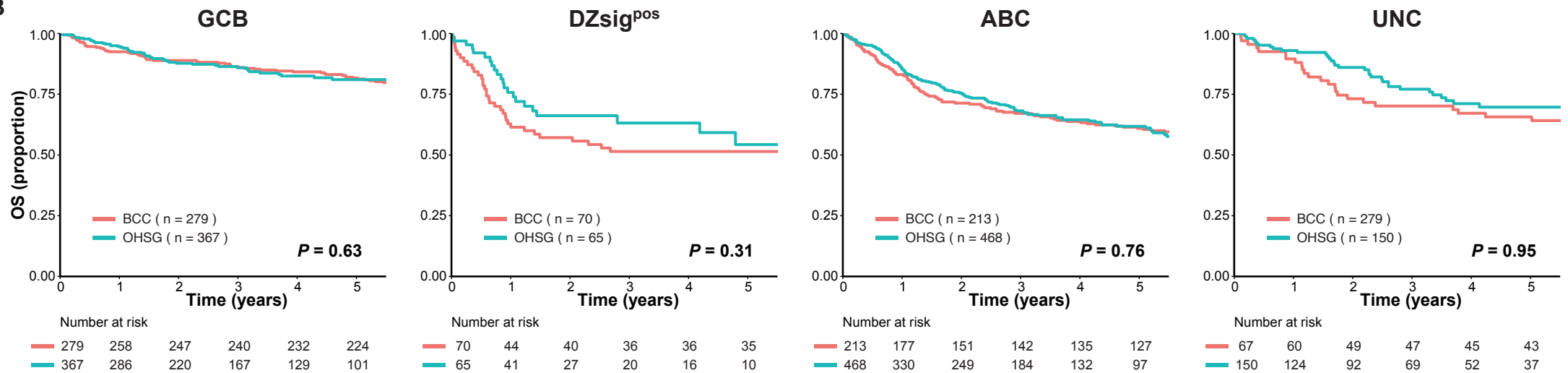
Supplemental Figure 4: Prognostic association of dark zone signature (DZsig) within each cell-of-origin (COO). Kaplan–Meier curves showing progression-free survival (PFS) and overall survival (OS) according to DZsig in patients with germinal center B-cell-like (GCB) tumors (A, B) and unclassified (UNC) tumors (C,D). ind: indeterminate, neg: negative, pos: positive

Supplemental Figure 7

A

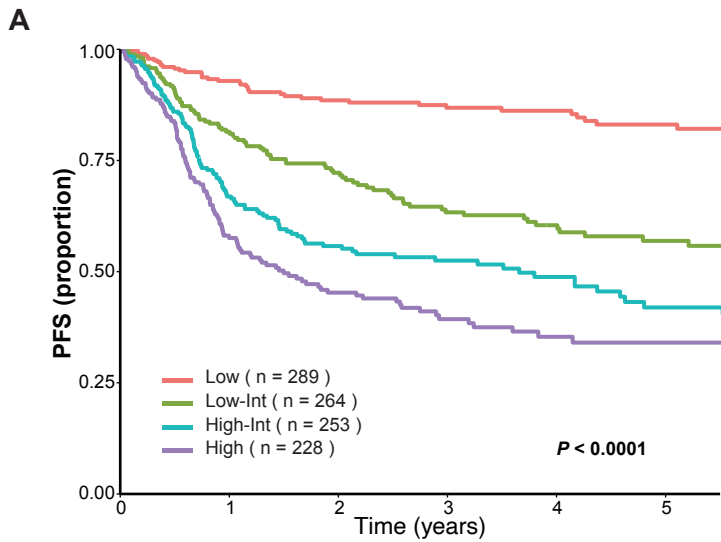


B

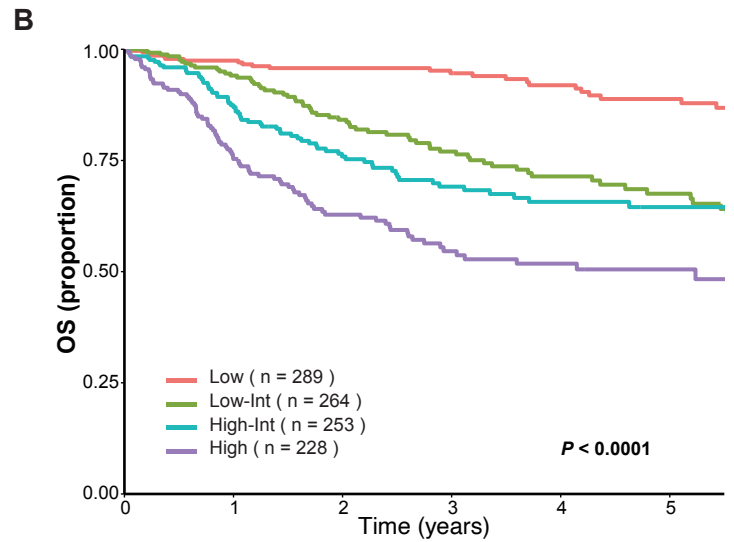


Supplemental Figure 7: Comparison of survival outcomes with each refined cell-of-origin according to cohorts. ABC: activated B-cell-like, BCC: BC Cancer, DZsig: dark zone signature, GCB: germinal center B-cell-like, OHSG: Okayama Hematology Study Group, pos: positive, UNC: unclassified

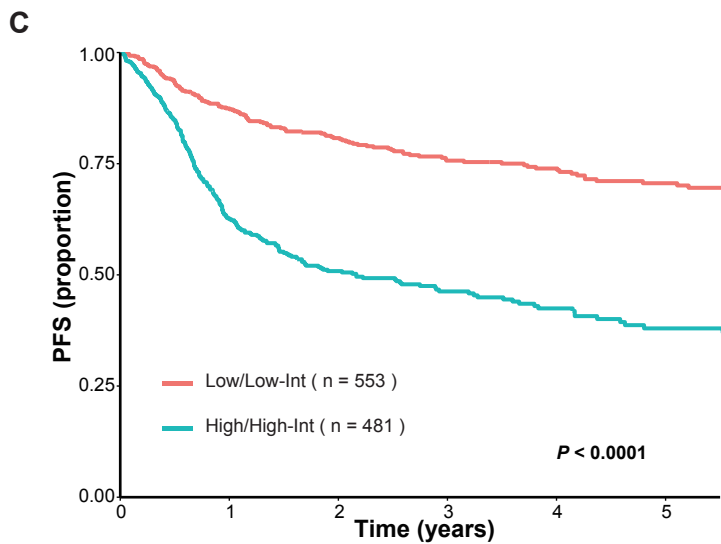
Supplemental Figure 8



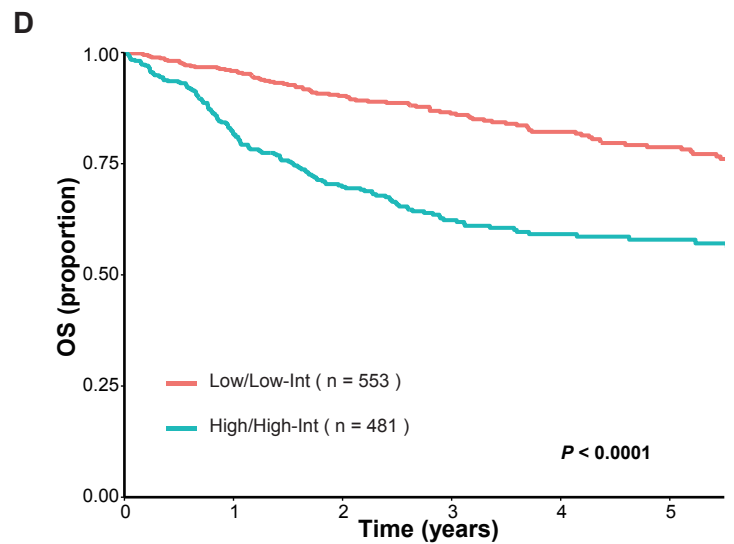
Number at risk						
	0	1	2	3	4	5
— Low (n = 289)	289	228	182	145	116	89
— Low-Int (n = 264)	264	177	135	99	76	54
— High-Int (n = 253)	253	144	97	67	48	33
— High (n = 228)	228	107	71	44	27	15



Number at risk						
	0	1	2	3	4	5
— Low (n = 289)	289	238	197	160	124	96
— Low-Int (n = 264)	264	204	157	121	88	64
— High-Int (n = 253)	253	187	126	89	66	52
— High (n = 228)	228	139	97	60	42	24



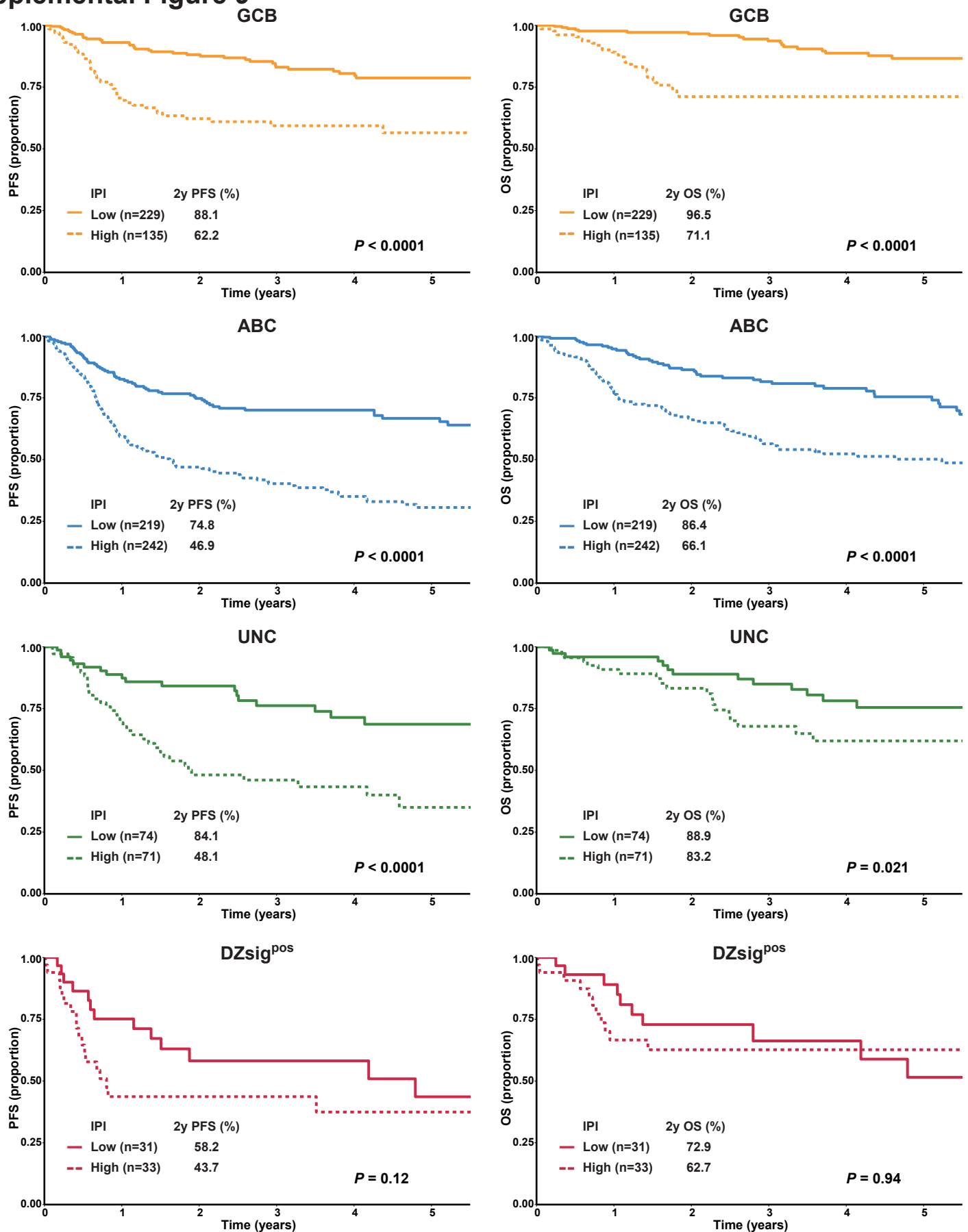
Number at risk						
	0	1	2	3	4	5
— Low/Low-Int (n = 553)	553	405	317	244	192	143
— High/High-Int (n = 481)	481	251	168	111	75	48



Number at risk						
	0	1	2	3	4	5
— Low/Low-Int (n = 553)	553	442	354	281	212	160
— High/High-Int (n = 481)	481	326	223	149	108	76

Supplemental Figure 8: Progression free survival (PFS) and overall survival (OS) stratified by international prognostic index (IPI). Kaplan–Meier curves showing PFS and OS according to each IPI group (A, B) and two IPI risk groups (C, D): low/low-intermediate risk IPI group and high/high-intermediate risk IPI group. Int: intermediate

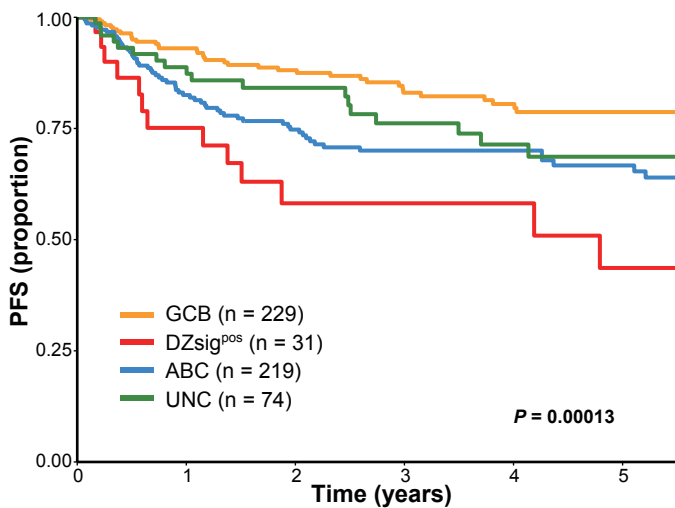
Supplemental Figure 9



Supplemental Figure 9: Progression free survival (PFS) and overall survival (OS) in patients with each refined cell-of-origin stratified by international prognostic index (IPI). Kaplan–Meier curves showing PFS and OS according to IPI risk groups: low/low-intermediate risk IPI group (Low) and high/high-intermediate risk IPI group (High). ABC: activated B-cell-like, DZsig: dark zone signature, GCB: germinal center B-cell-like, pos: positive, UNC: unclassified

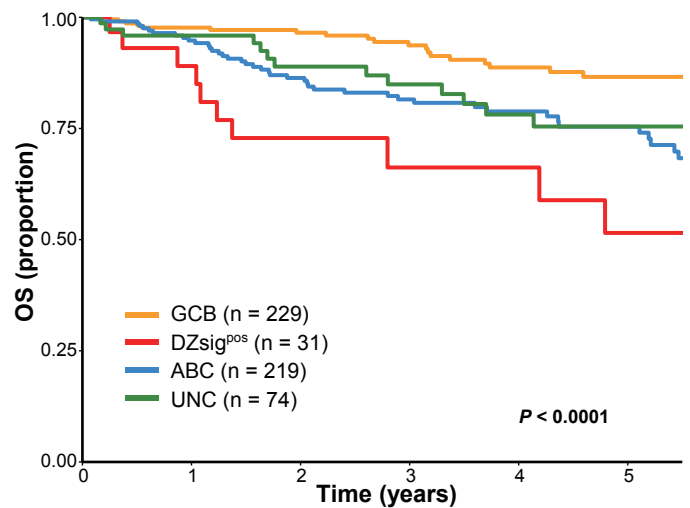
Supplemental Figure 10

A



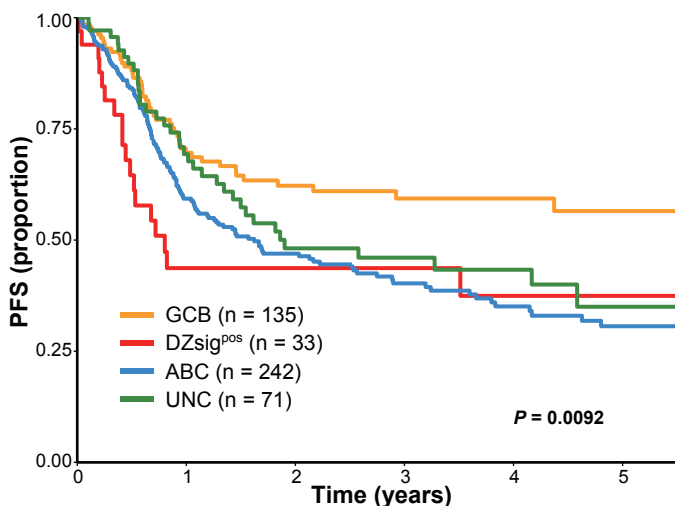
Number at risk						
	0	1	2	3	4	5
GCB (n = 229)	229	181	144	106	89	69
DZsig ^{pos} (n = 31)	31	19	12	9	8	5
ABC (n = 219)	219	145	116	92	68	50
UNC (n = 74)	74	60	45	37	27	19

B



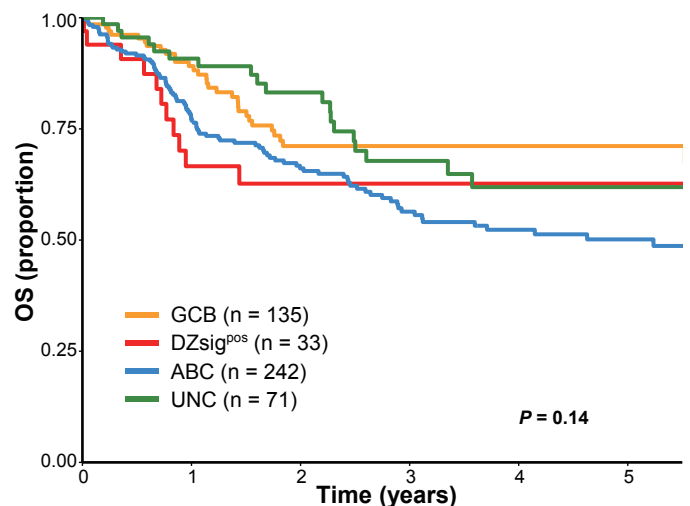
Number at risk						
	0	1	2	3	4	5
GCB (n = 229)	229	189	158	122	97	76
DZsig ^{pos} (n = 31)	31	22	14	10	9	6
ABC (n = 219)	219	166	134	107	76	56
UNC (n = 74)	74	65	48	42	30	22

C



Number at risk						
	0	1	2	3	4	5
GCB (n = 135)	135	73	52	35	22	17
DZsig ^{pos} (n = 33)	33	12	9	7	5	2
ABC (n = 242)	242	123	82	52	35	22
UNC (n = 71)	71	43	25	17	13	7

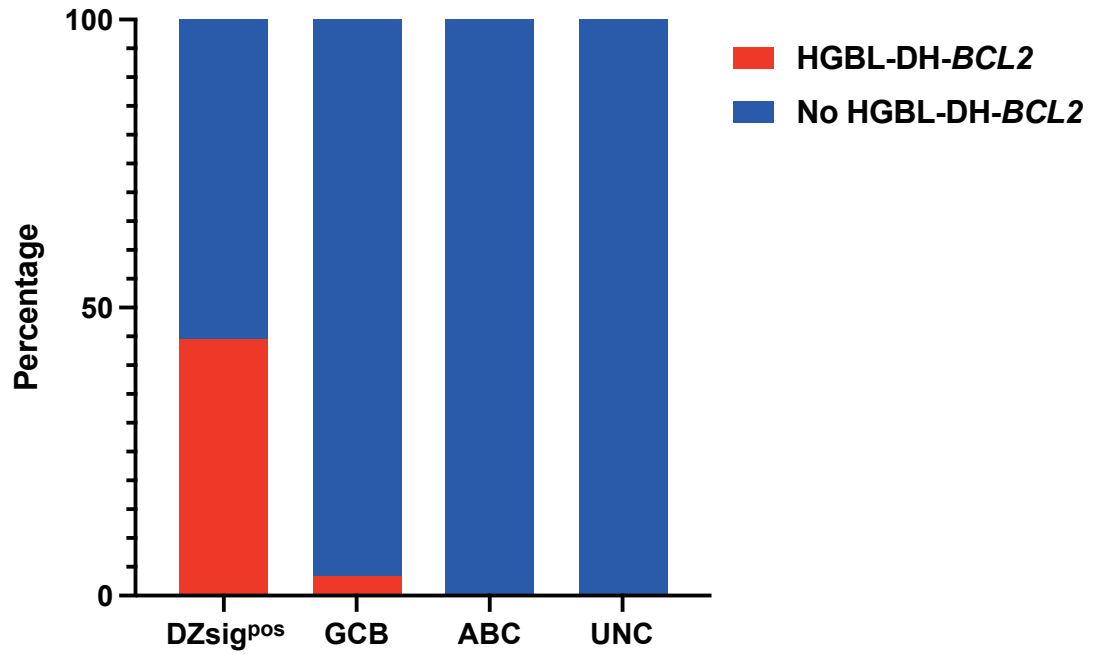
D



Number at risk						
	0	1	2	3	4	5
GCB (n = 135)	135	94	59	43	30	23
DZsig ^{pos} (n = 33)	33	19	13	10	7	4
ABC (n = 242)	242	158	111	73	53	38
UNC (n = 71)	71	55	40	23	18	11

Supplemental Figure 10: Prognostic association of refined cell-of-origin stratified by international prognostic index (IPI). (A, B) Kaplan–Meier curves showing PFS (A) and OS (B) for the low/low-intermediate risk IPI group. (C, D) Kaplan–Meier curves showing PFS (C) and OS (D) for the high/high-intermediate risk IPI group. ABC: activated B-cell-like, DZsig: dark zone signature, GCB: germinal center B-cell-like, pos: positive, UNC: unclassified

Supplemental Figure 11



Supplemental Figure 11: The proportion of high-grade B-cell lymphoma with *MYC* and *BCL2* rearrangements (HGBL-DH-*BCL2*) in each class. ABC: activated B-cell-like, GCB: germinal center B-cell-like, DZsig^{pos}: dark zone signature positive, UNC: unclassified