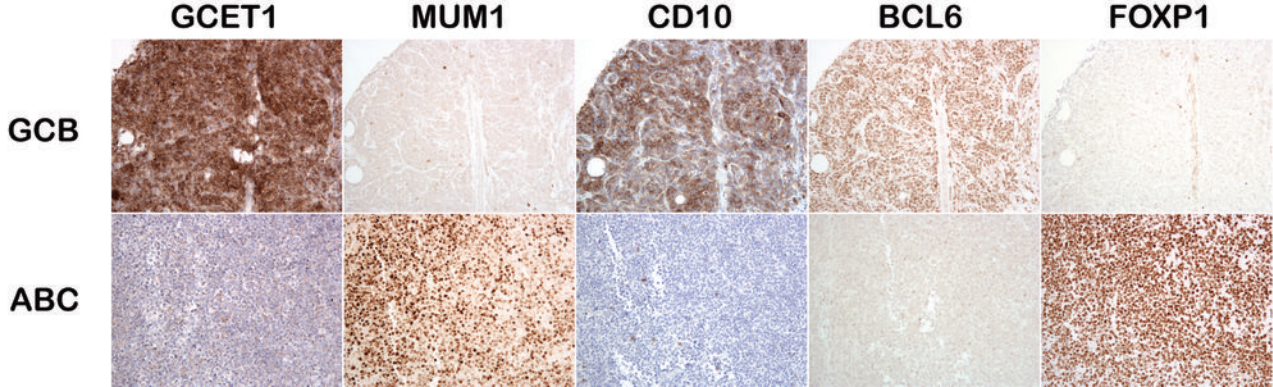
Supplemental Figure



Supplementary Table 1. Immunohistochemical stain protocols.

Antibody	Clone	Source	Dilution	Antigen Retrieval		
				Solution	Temperature (°C)	Duration (minutes)
GCET1	RAM341B/C1/C2*	Montes-Moreno *et al* (2008)[24]	Neat	1mM EDTA (pH 8.0)	115	10
MUM1	MUM1p†	Falini *et al* (2000)[26]	1:10	Cell Conditioning 1 (Ventana) (pH 8.0)	95	30
CD10	56C6	Ventana Medical Systems, Inc.	Neat	Cell Conditioning 1 (Ventana) (pH 8.0)	95	60
BCL6	PG-B6p	Dako A/S	1:60	Cell Conditioning 2 (Ventana) (pH 6.0)	95	30
FOXP1	JC12‡	Banham *et al* (2001)[27]	1:80	1mM EDTA (pH 8.0)	95	30
MTA3	polyclonal	Fujita *et al* (2003)[25]	1:1200	0.1M Tris (pH 10.0)	120	60
BCL6	polyclonal	Santa Cruz Biotechnology	1:75	1 mM EDTA (pH 8.0)	95	60
BCL2	124	Dako A/S	1:20	Cell Conditioning 1 (Ventana) (pH 8.0)	95	30
Cyclin D2	polyclonal	Santa Cruz Biotechnology	1:500	1 mM EDTA (pH 8.0)	95	60
CD20	L26	Dako A/S	1:200	Cell Conditioning 1 (Ventana) (pH 8.0)	95	8
CD3	PS1	Ventana Medical Systems, Inc.	Neat	Cell Conditioning 1 (Ventana) (pH 8.0)	95	30

EDTA indicates ethylenediaminetetraacetic acid.
*A commercial antibody for this clone is now available from Abcam Inc., Cambridge, MA, USA
†A commercial antibody for this clone is now available from Dako A/S, Carpinteria, CA, USA.
‡A commercial antibody for this clone is now available from Abcam Inc., Cambridge, MA, USA, and Novus Biologicals, Inc., Littleton, CO, USA.

Supplementary Table 2. Patient characteristics of the training set by gene expression profiling classification.

		Number (%)	GCB (%)	ABC (%)	P-value
		84	47	37	
Gender	Male	46 (55)	24 (51)	22 (59)	
	Female	38 (45)	23 (49)	15 (41)	0.440
Age (years)	< 60	33 (39)	21 (45)	12 (32)	
	≥ 60	51 (61)	26 (55)	25 (68)	0.250
Karnofsky score	> 70	69 (82)	43 (91)	26 (70)	
	≤ 70	15 (18)	4 (9)	11 (30)	0.012
Stage	I/II	37 (45)	23 (50)	14 (38)	
	III/IV	46 (55)	23 (50)	23 (62)	0.270
Extranodal sites	< 2	69 (84)	36 (77)	33 (94)	
	≥ 2	13 (16)	11 (24)	2 (6)	0.030
LDH	Normal	33 (45)	23 (53)	10 (33)	
	Elevated	40 (55)	20 (47)	20 (67)	0.089
IPI score	Low (0-2)	47 (67)	31 (74)	16 (57)	
	High (3-5)	23 (33)	11 (26)	12 (43)	0.150

LDH indicates lactate dehydrogenase; IPI, International Prognostic Index

Supplementary Table 3. Clinical characteristics of the validation set by gene expression profiling classification.

		Number (%)	GCB (%)	ABC (%)	P-value
		63	37	26	
Gender	Male	29 (46)	15 (41)	14 (54)	
	Female	34 (54)	22 (59)	12 (46)	0.300
Age (years)	< 60	29 (46)	21 (57)	8 (31)	
	≥ 60	34 (54)	16 (43)	18 (69)	0.071
Karnofsky	> 70	49 (83)	32 (91)	17 (71)	
Score	≤ 70	10 (17)	3 (9)	7 (29)	0.074
Stage	I/II	30 (50)	21 (60)	9 (36)	
	III/IV	30 (50)	14 (40)	16 (64)	0.120
Extranodal	< 2	50 (88)	30 (91)	20 (83)	
sites	≥ 2	7 (12)	3 (9)	4 (17)	0.440
LDH	Normal	19 (40)	14 (48)	5 (26)	
	Elevated	29 (60)	15 (52)	14 (74)	0.130
IPI score	Low (0-2)	27 (61)	20 (74)	7 (41)	
	High (3-5)	17 (39)	7 (26)	10 (59)	0.055

LDH indicates lactate dehydrogenase; IPI, International Prognostic Index

Supplementary Table 4. The percentages of positive cases for each immunostain in the GEP-defined subtypes of DLBCL. All positive results were defined as at least 30% of the lymphoma cells staining positive, except for GCET1, MUM1 and FOXP1, for which at least 80% of the positive lymphoma cells were required for a positive result.

	GCET1 (%)	CD10 (%)	Mono BCL6 (%)	Poly BCL6 (%)	MTA3 (%)	MUM1 (%)	FOXP1 (%)	BCL2 (%)	Cyclin D2 (%)
GCB (n=47)	57	57	89	83	67	6	30	50	0
ABC (n=37)	8	5	62	35	57	53	73	59	32
Unclassified (GCB by the new algorithm) (n=7)	14	43	86	29	43	14	14	43	0
Unclassified (ABC by the new algorithm) (n=12)	0	0	25	25	25	0	25	25	17
PMBL (n=7)	43	14	71	43	29	14	0	29	0