Table S1. Total number of pathology reports reviewed, clinical consideration for lymphoma and unique patient cases included in the study. In order to preserve tissue for future clinical needs, cases were excluded from the final cohort if blocks were <1 cm^3.

Year	Total pathology reports reviewed	Pathology reports from cases considered for lymphoma	Unique patients with pathology reports considered for lymphoma
2018	18	18	13
2017	7045	304	167
2016	8551	319	203
2015	7702	333	185
2014	8293	264	188
2013	8400	198	143
2012	8531	218	134
2011	7000	194	132
2010	7942	215	142
2009	8080	201	109
2008	7691	267	135
2007	7025	245	139
2006	7250	239	146
Total	**93528**	**3015**	**1836**

Table S2. A. WHO classification of 645 cases with corresponding chemical ligation-dependent probe amplification (CLPA) diagnostic bin. Patients whose samples failed the CLPA assay are also included. **B.** 32 additional malignant cases that passed QC and had a definitive WHO classification but were excluded from modeling due to low numbers for each entity with distinct biology.

A

WHO Classification	CLPA Diagnostic Bin	N
B lymphoblastic leukemia/lymphoma	Agg BCL	5
Burkitt lymphoma	Agg BCL	5
Diffuse large B cell lymphoma, germinal center type	DLBCL	68
Diffuse large B cell lymphoma, non-germinal center type	DLBCL	152
Diffuse large B cell lymphoma, NOS	DLBCL	7
EBV positive diffuse large B cell lymphoma, NOS	DLBCL	5
High grade B cell lymphoma, NOS	DLBCL	15
High grade B cell lymphoma, Myc failed	DLBCL	2
High grade B cell lymphoma with Myc and BCL2 or BCL6	DLBCL	9
T cell histocyte rich large B cell lymphoma	DLBCL	5
Follicular lymphoma, grade 1-2	FL	40
Follicular lymphoma, grade 3	FL	11
Classic Hodgkin lymphoma	HL	91
Nodular lymphocyte-predominant Hodgkin lymphoma	HL	6
Mantle cell lymphoma	MCL	63
Marginal zone lymphoma	MZL	23
Extranodal NK/T cell lymphoma, nasal type	NKTCL	54
Anaplastic large cell lymphoma, ALK-negative	TCL	7
Anaplastic large cell lymphoma, ALK-positive	TCL	6
Angioimmunoblastic T cell lymphoma	TCL	1
CD30 positive lymphoproliferative disorder	TCL	2
Mycosis Fungoides	TCL	1
Nodal peripheral T cell lymphoma with T follicular helper phenotype	TCL	2
Peripheral T cell lymphoma, NOS	TCL	19
Primary cutaneous CD4+ small/medium T-cell lymphoproliferative disorder	TCL	1
Nonmalignant cases		
Acute inflammation with ulceration	Nonmalignant	1
Adipose tissue	Nonmalignant	1
Angiomyomatous hamartoma	Nonmalignant	1

Chronic gastritis	Nonmalignant	2
Chronic inflammation	Nonmalignant	1
Dermatitis	Nonmalignant	1
Fibroconnective tissue	Nonmalignant	2
Fragmented lymphoid tissue	Nonmalignant	1
Granulomatous inflammation	Nonmalignant	1
Non-lesional tissue identified	Nonmalignant	1
Progressive transformation of germinal centers	Nonmalignant	1
Reactive follicular hyperplasia	Nonmalignant	29
Reactive lymphoid infiltrate	Nonmalignant	1
Reactive lymphoid hyperplasia	Nonmalignant	1
Normal skin	Nonmalignant	1
	Total Diagnoses	**645**

B

WHO Classification	**CLPA Diagnostic Bin**	**N**
Blastic plasmacytoid dendritic cell neoplasm	Excluded	1
Carcinoma	Excluded	3
CLL/SLL	Excluded	14
Neuroectoderm	Excluded	1
Plasma cell neoplasm	Excluded	7
Plasmablastic lymphoma	Excluded	1
T-lymphoblastic lymphoma	Excluded	5
	Total Excluded	**32**

Table S3. Comparison of blinded diagnoses utilizing H&E of whole section followed by IHC on tissue microarray (TMA) versus H&E and IHC on whole sections (n=80 cases). Two cases were excluded because of inadequate tissue. Percent tumor cell involvement was estimated from whole sections.

Case	% Tumor Cells	TMA Diagnosis	Whole Slide Diagnosis	Concordant	Comment
E0010	95	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0023	80	DLBCL, GC type	DLBCL, GC type	Y	
E0029	1	PTCL, NOS	Nodular lymphocyte predominant Hodgkin lymphoma	N	Rare large CD20+ cells noted on IHC of whole section and not sampled in TMA
E0035	50	DLBCL, GC type	DLBCL, GC type	Y	
E0061	90	DLBCL, GC type	DLBCL, GC type	Y	
E0065	90	Mantle cell lymphoma	Mantle cell lymphoma	Y	
E0076	50	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0085	20	ALCL, ALK positive	ALCL, ALK-positive	Y	
E0100	90	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0107	90	Burkitt lymphoma	High grade B cell lymphoma, NOS (favor Burkitt lymphoma)	Y	
E0109	80	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0115	95	DLBCL, GC type	DLBCL, GC type	Y	
E0127	90	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0154	50	DLBCL, GC type	DLBCL, GC type	Y	
E0178	80	DLBCL, non-GC type	DLBCL arising from marginal zone lymphoma	Y	
E0186	95	Mantle cell lymphoma	Mantle cell lymphoma	Y	
E0199	30	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0227	80	Mantle cell lymphoma	Mantle cell lymphoma	Y	
E0236	90	DLBCL, GC type	DLBCL, GC type (20%) arising from grade 3 FL (80%)	Y	
E0260	80	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0275	20	DLBCL, non-GC type	Limited tissue	N/A	Limited tissue on whole section IHC
E0317	N/A	Non-lymphoid tissue	Non-lymphoid tissue	Y	
E0322	25	DLBCL, non-GC type	DLBCL, GC type	Y	
E0334	1	Classical Hodgkin Lymphoma	Classical Hodgkin lymphoma	Y	
E0337	60	Follicular lymphoma, grade 3	Follicular lymphoma, grade 3 (10%); follicular lymphoma, grade 1-2 (90%)	Y	
E0343	80	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0378	90	Mantle cell lymphoma	Mantle cell lymphoma	Y	
E0380	90	Marginal zone lymphoma	Marginal zone lymphoma	Y	

E0381	1	Nodular lymphocyte predominant Hodgkin lymphoma with variant pattern	T cell histiocyte rich large B cell lymphoma arising from NLPHL	Y	On TMA, there is one large B cell nodule, so this is probably TCHRLB arising from NLPHL; this represents a biological continuum
E0397	80	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0442	40	DLBCL, non-GC type	DLBCL, non-GC type (80%) evolving from grade 3B FL (2%)	Y	
E0444	80	DLBCL, GC type	DLBCL, GC type (80%) evolving from grade 3B FL (2%)	Y	
E0511	90	Plasma cell neoplasm	Pleomorphic plasma cell neoplasm	Y	
E0522	50	ENKTL, nasal type	ENKTL, nasal type	Y	
E0525	1	Classical Hodgkin Lymphoma	Classical Hodgkin lymphoma	Y	
E0564	2	Classical Hodgkin Lymphoma	Classical Hodgkin lymphoma	Y	
E0571	80	Mantle cell lymphoma	Mantle cell lymphoma	Y	
E0597	95	Burkitt lymphoma	High grade B cell lymphoma, NOS (favor Burkitt lymphoma)	Y	
E0601	60	PTCL, NOS	PTCL/CD30+ lymphoproliferative disorder	Y	
E0602	2	Classical Hodgkin lymphoma	Classical Hodgkin lymphoma	Y	
E0604	90	Follicular lymphoma, grade 1-2	Follicular lymphoma, grade 1-2	Y	
E0609	90	DLBCL, GC type	DLBCL, GC type	Y	
E0610	70	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0612	80	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0641	20	Follicular lymphoma, grade 1-2	CD10+ small B cell lymphoma, favor follicular lymphoma	Y	
E0650	5	Classical Hodgkin lymphoma	Classical Hodgkin lymphoma	Y	
E0658	5	Classical Hodgkin Lymphoma	Classical Hodgkin lymphoma	Y	
E0679	90	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0709	80	DLBCL, GC type	DLBCL, GC type (10%) evolving from grade 3 FL (90%)	Y	
E0725	20	Classical Hodgkin Lymphoma	Classical Hodgkin lymphoma	Y	
E0736	80	Follicular lymphoma, grade 1-2	Follicular lymphoma, grade 1-2	Y	
E0743	50	Marginal zone lymphoma	Marginal zone lymphoma	Y	
E0744	80	Follicular lymphoma, grade 3	Follicular lymphoma, grade 3A	Y	
E0748	20	PTCL, NOS	PTCL, NOS	Y	
E0749	70	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0770	80	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0776	90	Mantle cell lymphoma	Mantle cell lymphoma	Y	
E0803	80	DLBCL, non-GC type	DLBCL, non-GC type	Y	

E0814	80	Mantle cell lymphoma	Mantle cell lymphoma	Y	
E0821	30	ALCL, ALK positive	ALCL, ALK-positive	Y	
E0829	90	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0853	80	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0872	90	DLBCL, GC type by IHC. FISH showed *MYC* and *BCL2* rearrangements	DLBCL, GC type (FISH unsuccessful)	Y	
E0877	90	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0886	70	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0928	70	DLBCL, non-GC type	DLBCL, non-GC type	Y	
E0930	70	Mantle cell lymphoma	Mantle cell lymphoma	Y	
E0939	0	Reactive follicular hyperplasia	Reactive follicular hyperplasia	Y	
E0940	0	Reactive follicular hyperplasia	Reactive follicular hyperplasia	Y	
E0941	0	Reactive follicular hyperplasia	Reactive follicular hyperplasia	Y	
E0946	0	Reactive follicular hyperplasia	Reactive follicular hyperplasia	Y	
E0953	0	Reactive follicular hyperplasia	Reactive follicular hyperplasia	Y	
E0962	10	ENKTL, nasal type	ENKTL, nasal type	Y	
E0964	N/A	PTCL, NOS	Non-diagnostic: scant tissue	N/A	No tissue to stain for IHC
E0974	60	ENKTL, nasal type	ENKTL, nasal type	Y	
E0975	80	ENKTL, nasal type	ENKTL, nasal type	Y	
E0980	5	ENKTL, nasal type	ENKTL, nasal type	Y	
E0987	60	ENKTL, nasal type	ENKTL, nasal type	Y	
E0991	60	ENKTL, nasal type	ENKTL, nasal type	Y	
E0992	N/A	Non-lesional tissue	Non- lesional tissue	Y	

Table S4. Characteristics of the 643 individual patients who underwent biopsy for suspected lymphoma in the pre-treatment or post-relapse setting. Patients whose samples failed the CLPA assay are also included. Two patients had both a pre-treatment and relapse tissue sample. Only the pre-treatment characteristics are included here.

	Agg BCL	DLBCL	FL	HL	MCL	MZL	NKTCL	Nonmalignant	TCL
N	10	262	50	97	63	23	54	45	39
Median age	38	63	59	38	62	53	41	48	56
(range)	(18-60)	(15-98)	(30-85)	(8-90)	(39-89)	(22-78)	(15-84)	(19-79)	(15-87)
Sex (%)									
Female	5 (50)	140 (53)	26 (52)	43 (44)	14 (22)	15 (65)	18 (34)	18 (40)	11 (28)
Male	5 (50)	119 (45)	21 (42)	53 (55)	49 (78)	8 (35)	32 (59)	7 (16)	27 (69)
Unknown	0 (0)	3 (1)	3 (6)	1 (1)	0 (0)	0 (0)	4 (7)	20 (44)	1 (3)
Stage (%)									
I	0 (0)	51 (19)	6 (12)	19 (20)	1 (2)	5 (22)	18 (33)	6 (13)	8 (21)
II	0 (0)	60 (23)	11 (22)	31 (32)	10 (16)	4 (17)	20 (37)	5 (11)	7 (18)
III	2 (20)	57 (22)	14 (28)	23 (24)	19 (30)	10 (43)	0 (0)	4 (9)	5 (13)
IV	5 (50)	41 (16)	6 (12)	9 (9)	25 (40)	2 (9)	5 (9)	0 (0)	9 (23)
Unknown	3 (30)	53 (20)	13 (26)	15 (15)	8 (13)	2 (9)	11 (20)	30 (67)	10 (26)
B symptoms (%)									
Yes	4 (40)	88 (34)	11 (22)	49 (51)	30 (48)	8 (35)	15 (28)	1 (2)	11 (28)
No	3 (30)	120 (46)	26 (52)	33 (34)	25 (40)	13 (57)	27 (50)	14 (31)	19 (49)
Unknown	3 (30)	54 (21)	13 (26)	15 (15)	8 (13)	2 (9)	12 (22)	30 (67)	9 (23)
Bulky (%)									
Yes	3 (30)	113 (43)	15 (30)	47 (48)	30 (48)	3 (13)	16 (30)	2 (4)	13 (33)
No	4 (40)	94 (36)	22 (44)	33 (34)	23 (37)	18 (78)	18 (33)	12 (27)	16 (41)
Unknown	3 (30)	55 (21)	13 (26)	17 (18)	10 (16)	2 (9)	20 (37)	31 (69)	10 (26)

Table S5. Individual diagnostic, clinical characteristics and treatment. Diagnosis at INCAN was obtained based on H&E staining with or without immunohistochemistry (IHC), as indicated in column C. Stanford diagnoses were based on the 2016 WHO classification as outlined in the Methods. Column E compares the diagnosis between INCAN and Stanford. "FULL" concordance indicates an identical finalized diagnosis. "PARTIAL" concordance indicates cases where the INCAN diagnosis and the Stanford diagnosis were within similar but not identical morphologic groupings. Any case in which the cell lineage between B, T, and NK cell lymphoma were discordant were considered completely discordant and indicated by "NONE." Lymphoma bin indicates the diagnostic bin for each case based on the Stanford diagnosis. Chemical ligation-dependent probe amplification (CLPA) diagnosis indicates the highest probability bin as indicated by the CLPA assay. In Column H, cases are classified as "indeterminate" if the CLPA probability call was <60%, "Yes" if the CLPA diagnosis differed from the INCAN diagnosis, and "No" if the CLPA diagnosis was identical to the INCAN diagnosis. Departments in Guatemala (Column M) are synonymous with states or provinces. Stage at diagnosis is based on review of the medical record and imaging but patients may have been understaged based on incomplete imaging or bone marrow assessment. HIV, Hepatitis B (HBsAb) and Hepatitis C (HCV Ab) status were based on antibody testing. The prognostic index was determined based upon the patient's original diagnosis and appropriate risk scoring system (IPI for DLBCL, PTCL, NKTCL, and MZL, MIPI for MCL, FLIPI for FL, Hasenclever for HL). Abbreviations: follicular lymphoma (FL), diffuse large B cell lymphoma (DLBCL), peripheral T cell lymphoma (PTCL), mantle cell lymphoma (MCL), aggressive B-cell lymphoma (AGG_BCL), Hodgkin lymphoma (HL), extranodal NK/T cell lymphoma (NKTCL), anaplastic large cell lymphoma (ALCL), cyclophosphamide/daunorubicin/vincristine/prednisone (CHOP), doxorubicin/bleomycin/vinblastine/dacarbazine (ABVD), gemcitabine/cyclophosphamide/vincristine/prednisone (GCVP), cyclophosphamide/vincristine/prednisone (CVP), dexamethasone/high dose cytarabine/cisplatin (DHAP), dexamethasone/methotrexate/ifosfamide/L-asparaginase/etoposide (SMILE) and rituximab (R).

See attached .xls file

Table S6. Concordance for 643 patient samples obtained for suspicion of lymphoma and assessed at both INCAN and Stanford. At INCAN, samples were evaluated by hematoxylin and eosin (H&E) staining alone or H&E with IHC. At Stanford, all samples were evaluated using H&E and IHC, as well as FISH in a subset. Gray boxes are considered concordant. In some cases, it was not possible to ensure concordance; for example, high-grade BCL at INCAN does not distinguish between lymphomas like Burkitt or LBL (which were included in the Agg BCL bin) and DLBCL. Overall, 244 (38%; 95% CI: 34-42%) were concordant. Additional abbreviations: CLL/SLL, chronic lymphocytic leukemia/small lymphocytic lymphoma; INC, incomplete diagnosis; T-LL, T-cell lymphoblastic lymphoma.

INCAN	**Stanford**									
Frequency	AGG BCL	DLBCL	FL	HL	MCL	MZL	NK-TCL	Nonmalignant	TCL	Total
AGG BCL	2	3	0	0	0	0	1	0	1	7
CLL/SLL	0	1	0	0	0	1	0	0	0	2
DLBCL	1	72	3	0	1	1	0	1	2	81
FL	1	7	20	0	6	0	0	0	0	34
HL	0	3	0	65	0	1	0	1	4	74
INC	5	160	27	16	35	13	12	31	19	318
MCL	0	1	0	0	20	0	0	0	0	21
MZL	0	3	0	0	0	6	0	2	0	11
NK-TCL	0	0	0	1	0	0	38	1	0	40
Nonmalignant	0	3	0	2	0	1	0	8	0	14
T-LL	1	0	0	0	0	0	0	0	0	1
TCL	0	9	0	13	1	0	3	1	13	40
Total	10	262	50	97	63	23	54	45	39	643

Table S7. Genes included in the CLPA-based assay and selected publications supporting their inclusion. Abbreviations and references are noted specifically below.

***Gene*/Protein**	**Gene Expression Discriminations**
FOXP1	ABC-DLBCL vs GCB-DLBCL[1-4] BL and molecular high-grade vs GCB-DLBCL[5]
TNFRSF13B/ CD267/TACI	ABC-DLBCL vs GCB-DLBCL[6,7]and PMBCL[2] MYC+/BCL2+ vs MYC-/BCL2- DLBCL[8] NMZL vs FL and reactive lymph node[9]
IGHM	ABC-DLBCL vs GCB-DLBCL[1,6] and PMBCL[2]
IRF4/MUM1	ABC-DLBCL and PMBCL[2] vs GCB-DLBCL[1,3,7,10] DLBCL vs Burkitts[11] NMZL vs FL and reactive lymph node[9]
BCL6	GCB-DLBC vs ABC-DLBCL[1,3,6,8] FL vs NMZL[9]. AITL vs ALCL, ATLL, NKTCL, PTC-NOS[12] ALK+ vs ALK- ALCL[13]
NEK6	GCB-DLBCL and PMBCL[2] vs ABC DLBCL[6] AITL vs ALCL, ATLL, NKTCL, PTC-NOS[12]
LMO2	GCB-DLBCL and PMBCL vs ABC-DLBCL[1-3,6,10,14,15] and BL[11] FL and reactive lymph node vs NMZL[9] Reactive lymph node, cHL and DLBCL vs FL[16]
MYBL1/AMYB	GCB-DLBCL, vs ABC-DLBCL[1,3,6,8] and PMBCL[2] BL vs DLBCL[11]
ALK	ALK+ expression improves prognosis vs ALK- in ALCL and is therapeutic target[13,17] Poor prognosis in DLBCL[18]
BCL2	DLBCL vs BL[11,19] DLBCL vs PMBCL[20] AITL vs ALCL, ATLL, NKTCL/T, PTCL-NOS[12]
BCL2A1/BFL1	PMBCL vs HL and DLBCL[20] DLBCL vs BL[11,19,21,22] AITL vs ALCL, ATLL, NKTCL, PTCL-NOS[12]
BMP7	BL vs DLBCL[11,19]
CCND1	MCL vs SLL and MZL[23]
CD244/NKR2B4	NKTCL vs PTCL-NOS[24]

	NKTCL vs AITL, ATLL, ALCL[25] CT-PTCL vs ALK+ ALCL, ATLL, AITL [12]
CD44	PMBCL, DLBCL vs BL[11,19,22] ABC-DLBCL vs GCB-DLBCL[10,20,26] NMZL vs FL [9]
DLEU1	BL vs DLBCL[19,22]
CD5	SLL and MCL vs NMZL and FL
CRBN	Potential biomarker for imides
EBER	NKTCL, DLBCL, HL, AITL and potential biomarker
FCER2/CD23	PMBCL vs DLBCL[2] SLL vs MCL, MZL AITL vs ALCL, ATLL, NK/T, PTCL-NOS[12]
GATA3	PTCL-NOS with adverse prognosis[25,27] HL (cell lines) vs normal B cells, BL, FL, CLL[28] ATLL vs ALK+ALCL, ATLL, AITL, CT-PTCL [12]
ICOS	AITL vs ALCL, ATLL, NKTCL, PTCL-NOS[12] PTCL vs NKTCL[29]
ID3	AITL vs ALCL, ATLL, NKTCL, PTC-NOS[12] BL vs DLBCL[11,21,30]
MAL	PMBCL vs DLBCL[2]
MKI67	MCL poor prognosis[31] Aggressive versus indolent
MME/CD10	GCB-DLBC vs ABC-DLBCL[10,26] and BL[2,4,11,19,21] FL vs NMZL[9]
MS4A1/CD20	Mature B-cell lymphomas
MYC/c-MYC	BL and poor prognosis DLBCL[2,8,11,19,32-34] NKTCL vs normal peripheral blood NK cells[24]
NCAM1/CD56	NKTCL vs PTCL[24]
NFKBIA	DLBCL vs BL[11,19,21,30]
PAX5	B-cell malignancies
REL/c-Rel	PMBCL versus DLBCL[4,14]
SOX8	AITL vs ALCL, ATLL, NKTCL, PTCL-NOS[12]
STAT3	DLBCL vs BL[11,21,30]
TBX21/TBET	NKTCL vs PTCL[24,29] PTCL-NOS with better prognosis[25]

TCF3/E2A	BL and molecular high-grade vs DLBCL[19,21,30]
TNFRSF8/CD30	ALCL and subset of other PTCL and CTCL[12,25,29] PMBCL vs DLBCL[2]

Abbreviations:

ABC	activated B-cell
AITL	angioimmunoblastic t-cell lymphoma
ALCL	anaplastic large-cell lymphoma
ALK	anaplastic lymphoma kinase
ATLL	adult t-cell leukemia/lymphoma
BL	Burkitt lymphoma
cHL	classical Hodgkin lymphoma
CT-PTCL	cytotoxic peripheral t-cell lymphoma
CTCL	cutaneous t-cell lymphoma
DLBCL	diffuse large B-Cell Lymphoma
FL	follicular lymphoma
GCB	germinal center B-cell
HL	Hodgkin lymphoma
MCL	mantle cell lymphoma
NKTCL	natural killer/ t-cell lymphoma
NMZL	nodal marginal zone b-cell lymphoma
NOS	not otherwise specified
PMBCL	primary mediastinal large B-cell lymphoma
PTCL	peripheral t-cell lymphoma
SLL	small lymphocytic lymphoma/leukemia

References

1. Blenk S, Engelmann J, Weniger M, Schultz J, Dittrich M, Rosenwald A, Muller-Hermelink HK, Muller T, Dandekar T. Germinal center B cell-like (GCB) and activated B cell-like (ABC) type of diffuse large B cell lymphoma (DLBCL): analysis of molecular predictors, signatures, cell cycle state and patient survival. *Cancer Inform*. 2007;3:399-420.
2. Bobee V, Ruminy P, Marchand V, Viailly PJ, Abdel Sater A, Veresezan L, Drieux F, Berard C, Bohers E, Mareschal S, Dubois S, Jais JP, Leroy K, Figeac M, Picquenot JM, Molina TJ, Salles G, Haioun C, Tilly H, Jardin F. Determination of Molecular Subtypes of Diffuse Large B-Cell Lymphoma Using a Reverse Transcriptase Multiplex Ligation-Dependent Probe Amplification Classifier: A CALYM Study. *J Mol Diagn.* 2017;19(6):892-904.
3. Jais JP, Haioun C, Molina TJ, Rickman DS, de Reynies A, Berger F, Gisselbrecht C, Briere J, Reyes F, Gaulard P, Feugier P, Labouyrie E, Tilly H, Bastard C, Coiffier B, Salles G, Leroy K, Groupe d'Etude des Lymphomes de lA. The expression of 16 genes related to the cell of

origin and immune response predicts survival in elderly patients with diffuse large B-cell lymphoma treated with CHOP and rituximab. *Leukemia*. 2008;22(10):1917-1924.
4. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A*. 2003;100(17):9991-9996.
5. Sha C, Barrans S, Cucco F, Bentley MA, Care MA, Cummin T, Kennedy H, Thompson JS, Uddin R, Worrillow L, Chalkley R, van Hoppe M, Ahmed S, Maishman T, Caddy J, Schuh A, Mamot C, Burton C, Tooze R, Davies A, Du MQ, Johnson PWM, Westhead DR. Molecular High-Grade B-Cell Lymphoma: Defining a Poor-Risk Group That Requires Different Approaches to Therapy. *J Clin Oncol*. 2019;37(3):202-212.
6. Mareschal S, Ruminy P, Bagacean C, Marchand V, Cornic M, Jais JP, Figeac M, Picquenot JM, Molina TJ, Fest T, Salles G, Haioun C, Leroy K, Tilly H, Jardin F. Accurate Classification of Germinal Center B-Cell-Like/Activated B-Cell-Like Diffuse Large B-Cell Lymphoma Using a Simple and Rapid Reverse Transcriptase-Multiplex Ligation-Dependent Probe Amplification Assay: A CALYM Study. *J Mol Diagn*. 2015.
7. Scott DW, Wright GW, Williams PM, Lih CJ, Walsh W, Jaffe ES, Rosenwald A, Campo E, Chan WC, Connors JM, Smeland EB, Mottok A, Braziel RM, Ott G, Delabie J, Tubbs RR, Cook JR, Weisenburger DD, Greiner TC, Glinsmann-Gibson BJ, Fu K, Staudt LM, Gascoyne RD, Rimsza LM. Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin-embedded tissue. *Blood*. 2014;123(8):1214-1217.
8. Hu S, Xu-Monette ZY, Tzankov A, Green T, Wu L, Balasubramanyam A, Liu WM, Visco C, Li Y, Miranda RN, Montes-Moreno S, Dybkaer K, Chiu A, Orazi A, Zu Y, Bhagat G, Richards KL, Hsi ED, Choi WW, Zhao X, van Krieken JH, Huang Q, Huh J, Ai W, Ponzoni M, Ferreri AJ, Zhou F, Slack GW, Gascoyne RD, Tu M, Variakojis D, Chen W, Go RS, Piris MA, Moller MB, Medeiros LJ, Young KH. MYC/BCL2 protein coexpression contributes to the inferior survival of activated B-cell subtype of diffuse large B-cell lymphoma and demonstrates high-risk gene expression signatures: a report from The International DLBCL Rituximab-CHOP Consortium Program. *Blood*. 2013;121(20):4021-4031; quiz 4250.
9. Arribas AJ, Campos-Martin Y, Gomez-Abad C, Algara P, Sanchez-Beato M, Rodriguez-Pinilla MS, Montes-Moreno S, Martinez N, Alves-Ferreira J, Piris MA, Mollejo M. Nodal marginal zone lymphoma: gene expression and miRNA profiling identify diagnostic markers and potential therapeutic targets. *Blood*. 2012;119(3):e9-e21.
10. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J, Jr., Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. 2000;403(6769):503-511.
11. Dave SS, Fu K, Wright GW, Lam LT, Kluin P, Boerma EJ, Greiner TC, Weisenburger DD, Rosenwald A, Ott G, Muller-Hermelink HK, Gascoyne RD, Delabie J, Rimsza LM, Braziel RM, Grogan TM, Campo E, Jaffe ES, Dave BJ, Sanger W, Bast M, Vose JM, Armitage JO, Connors JM, Smeland EB, Kvaloy S, Holte H, Fisher RI, Miller TP, Montserrat E, Wilson WH, Bahl M, Zhao H, Yang L, Powell J, Simon R, Chan WC, Staudt LM, Lymphoma/Leukemia Molecular Profiling P. Molecular diagnosis of Burkitt's lymphoma. *N Engl J Med*. 2006;354(23):2431-2442.
12. Iqbal J, Weisenburger DD, Greiner TC, Vose JM, McKeithan T, Kucuk C, Geng H, Deffenbacher K, Smith L, Dybkaer K, Nakamura S, Seto M, Delabie J, Berger F, Loong F, Au WY, Ko YH, Sng I, Armitage JO, Chan WC, International Peripheral TCLP. Molecular signatures to improve diagnosis in peripheral T-cell lymphoma and prognostication in angioimmunoblastic T-cell lymphoma. *Blood*. 2010;115(5):1026-1036.

13. Lamant L, de Reynies A, Duplantier MM, Rickman DS, Sabourdy F, Giuriato S, Brugieres L, Gaulard P, Espinos E, Delsol G. Gene-expression profiling of systemic anaplastic large-cell lymphoma reveals differences based on ALK status and two distinct morphologic ALK+ subtypes. *Blood*. 2007;109(5):2156-2164.
14. Rosenwald A, Wright G, Leroy K, Yu X, Gaulard P, Gascoyne RD, Chan WC, Zhao T, Haioun C, Greiner TC, Weisenburger DD, Lynch JC, Vose J, Armitage JO, Smeland EB, Kvaloy S, Holte H, Delabie J, Campo E, Montserrat E, Lopez-Guillermo A, Ott G, Muller-Hermelink HK, Connors JM, Braziel R, Grogan TM, Fisher RI, Miller TP, LeBlanc M, Chiorazzi M, Zhao H, Yang L, Powell J, Wilson WH, Jaffe ES, Simon R, Klausner RD, Staudt LM. Molecular diagnosis of primary mediastinal B cell lymphoma identifies a clinically favorable subgroup of diffuse large B cell lymphoma related to Hodgkin lymphoma. *J Exp Med*. 2003;198(6):851-862.
15. Dave SS, Fu K, Wright GW, Lam LT, Kluin P, Boerma EJ, Greiner TC, Weisenburger DD, Rosenwald A, Ott G, Muller-Hermelink HK, Gascoyne RD, Delabie J, Rimsza LM, Braziel RM, Grogan TM, Campo E, Jaffe ES, Dave BJ, Sanger W, Bast M, Vose JM, Armitage JO, Connors JM, Smeland EB, Kvaloy S, Holte H, Fisher RI, Miller TP, Montserrat E, Wilson WH, Bahl M, Zhao H, Yang L, Powell J, Simon R, Chan WC, Staudt LM. Molecular diagnosis of Burkitt's lymphoma. *N Engl J Med*. 2006;354(23):2431-2442.
16. Loi TH, Campain A, Bryant A, Molloy TJ, Lutherborrow M, Turner J, Yang YH, Ma DD. Discriminating lymphomas and reactive lymphadenopathy in lymph node biopsies by gene expression profiling. *BMC Med Genomics*. 2011;4:27.
17. Gascoyne RD, Aoun P, Wu D, Chhanabhai M, Skinnider BF, Greiner TC, Morris SW, Connors JM, Vose JM, Viswanatha DS, Coldman A, Weisenburger DD. Prognostic significance of anaplastic lymphoma kinase (ALK) protein expression in adults with anaplastic large cell lymphoma. *Blood*. 1999;93(11):3913-3921.
18. Laurent C, Do C, Gascoyne RD, Lamant L, Ysebaert L, Laurent G, Delsol G, Brousset P. Anaplastic lymphoma kinase-positive diffuse large B-cell lymphoma: a rare clinicopathologic entity with poor prognosis. *J Clin Oncol*. 2009;27(25):4211-4216.
19. Carey CD, Gusenleitner D, Chapuy B, Kovach AE, Kluk MJ, Sun HH, Crossland RE, Bacon CM, Rand V, Dal Cin P, Le LP, Neuberg D, Sohani AR, Shipp MA, Monti S, Rodig SJ. Molecular classification of MYC-driven B-cell lymphomas by targeted gene expression profiling of fixed biopsy specimens. *J Mol Diagn*. 2015;17(1):19-30.
20. Feuerhake F, Kutok JL, Monti S, Chen W, LaCasce AS, Cattoretti G, Kurtin P, Pinkus GS, de Leval L, Harris NL, Savage KJ, Neuberg D, Habermann TM, Dalla-Favera R, Golub TR, Aster JC, Shipp MA. NFkappaB activity, function, and target-gene signatures in primary mediastinal large B-cell lymphoma and diffuse large B-cell lymphoma subtypes. *Blood*. 2005;106(4):1392-1399.
21. Hummel M, Bentink S, Berger H, Klapper W, Wessendorf S, Barth TF, Bernd HW, Cogliatti SB, Dierlamm J, Feller AC, Hansmann ML, Haralambieva E, Harder L, Hasenclever D, Kuhn M, Lenze D, Lichter P, Martin-Subero JI, Moller P, Muller-Hermelink HK, Ott G, Parwaresch RM, Pott C, Rosenwald A, Rosolowski M, Schwaenen C, Sturzenhofecker B, Szczepanowski M, Trautmann H, Wacker HH, Spang R, Loeffler M, Trumper L, Stein H, Siebert R, Molecular Mechanisms in Malignant Lymphomas Network Project of the Deutsche K. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N Engl J Med*. 2006;354(23):2419-2430.
22. Morgan EA, Sweeney MP, Tomoka T, Kopp N, Gusenleitner D, Redd RA, Carey CD, Masamba L, Kamiza S, Pinkus GS, Neuberg DS, Rodig SJ, Milner DA, Jr., Weinstock DM. Targetable subsets of non-Hodgkin lymphoma in Malawi define therapeutic opportunities. *Blood Adv*. 2016;1(1):84-92.
23. Thieblemont C, Nasser V, Felman P, Leroy K, Gazzo S, Callet-Bauchu E, Loriod B, Granjeaud S, Gaulard P, Haioun C, Traverse-Glehen A, Baseggio L, Bertucci F, Birnbaum D, Magrangeas F, Minvielle S, Avet-Loiseau H, Salles G, Coiffier B, Berger F, Houlgatte R. Small

lymphocytic lymphoma, marginal zone B-cell lymphoma, and mantle cell lymphoma exhibit distinct gene-expression profiles allowing molecular diagnosis. *Blood*. 2004;103(7):2727-2737.
24. Huang Y, de Reynies A, de Leval L, Ghazi B, Martin-Garcia N, Travert M, Bosq J, Briere J, Petit B, Thomas E, Coppo P, Marafioti T, Emile JF, Delfau-Larue MH, Schmitt C, Gaulard P. Gene expression profiling identifies emerging oncogenic pathways operating in extranodal NK/T-cell lymphoma, nasal type. *Blood*. 2010;115(6):1226-1237.
25. Iqbal J, Wright G, Wang C, Rosenwald A, Gascoyne RD, Weisenburger DD, Greiner TC, Smith L, Guo S, Wilcox RA, Teh BT, Lim ST, Tan SY, Rimsza LM, Jaffe ES, Campo E, Martinez A, Delabie J, Braziel RM, Cook JR, Tubbs RR, Ott G, Geissinger E, Gaulard P, Piccaluga PP, Pileri SA, Au WY, Nakamura S, Seto M, Berger F, de Leval L, Connors JM, Armitage J, Vose J, Chan WC, Staudt LM, Lymphoma Leukemia Molecular Profiling P, the International Peripheral TcLP. Gene expression signatures delineate biological and prognostic subgroups in peripheral T-cell lymphoma. *Blood*. 2014;123(19):2915-2923.
26. Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, Gascoyne RD, Muller-Hermelink HK, Smeland EB, Giltnane JM, Hurt EM, Zhao H, Averett L, Yang L, Wilson WH, Jaffe ES, Simon R, Klausner RD, Powell J, Duffey PL, Longo DL, Greiner TC, Weisenburger DD, Sanger WG, Dave BJ, Lynch JC, Vose J, Armitage JO, Montserrat E, Lopez-Guillermo A, Grogan TM, Miller TP, LeBlanc M, Ott G, Kvaloy S, Delabie J, Holte H, Krajci P, Stokke T, Staudt LM, Lymphoma/Leukemia Molecular Profiling P. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med*. 2002;346(25):1937-1947.
27. Heavican TB, Bouska A, Yu J, Lone W, Amador C, Gong Q, Zhang W, Li Y, Dave BJ, Nairismagi ML, Greiner TC, Vose J, Weisenburger DD, Lachel C, Wang C, Fu K, Stevens JM, Lim ST, Ong CK, Gascoyne RD, Missiaglia E, Lemonnier F, Haioun C, Hartmann S, Pedersen MB, Laginestra MA, Wilcox RA, Teh BT, Yoshida N, Ohshima K, Seto M, Rosenwald A, Ott G, Campo E, Rimsza LM, Jaffe ES, Braziel RM, d'Amore F, Inghirami G, Bertoni F, de Leval L, Gaulard P, Staudt LM, McKeithan TW, Pileri S, Chan WC, Iqbal J. Genetic drivers of oncogenic pathways in molecular subgroups of peripheral T-cell lymphoma. *Blood*. 2019;133(15):1664-1676.
28. Kuppers R, Klein U, Schwering I, Distler V, Brauninger A, Cattoretti G, Tu Y, Stolovitzky GA, Califano A, Hansmann ML, Dalla-Favera R. Identification of Hodgkin and Reed-Sternberg cell-specific genes by gene expression profiling. *J Clin Invest*. 2003;111(4):529-537.
29. Iqbal J, Weisenburger DD, Chowdhury A, Tsai MY, Srivastava G, Greiner TC, Kucuk C, Deffenbacher K, Vose J, Smith L, Au WY, Nakamura S, Seto M, Delabie J, Berger F, Loong F, Ko YH, Sng I, Liu X, Loughran TP, Armitage J, Chan WC, International Peripheral TcLP. Natural killer cell lymphoma shares strikingly similar molecular features with a group of non-hepatosplenic gammadelta T-cell lymphoma and is highly sensitive to a novel aurora kinase A inhibitor in vitro. *Leukemia*. 2011;25(2):348-358.
30. Sha C, Barrans S, Care MA, Cunningham D, Tooze RM, Jack A, Westhead DR. Transferring genomics to the clinic: distinguishing Burkitt and diffuse large B cell lymphomas. *Genome Med*. 2015;7(1):64.
31. Scott DW, Abrisqueta P, Wright GW, Slack GW, Mottok A, Villa D, Jares P, Rauert-Wunderlich H, Royo C, Clot G, Pinyol M, Boyle M, Chan FC, Braziel RM, Chan WC, Weisenburger DD, Cook JR, Greiner TC, Fu K, Ott G, Delabie J, Smeland EB, Holte H, Jaffe ES, Steidl C, Connors JM, Gascoyne RD, Rosenwald A, Staudt LM, Campo E, Rimsza LM, Lymphoma/Leukemia Molecular Profiling P. New Molecular Assay for the Proliferation Signature in Mantle Cell Lymphoma Applicable to Formalin-Fixed Paraffin-Embedded Biopsies. *J Clin Oncol*. 2017;35(15):1668-1677.
32. Green TM, Young KH, Visco C, Xu-Monette ZY, Orazi A, Go RS, Nielsen O, Gadeberg OV, Mourits-Andersen T, Frederiksen M, Pedersen LM, Moller MB. Immunohistochemical double-hit score is a strong predictor of outcome in patients with diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. *J Clin Oncol*. 2012;30(28):3460-3467.

33. Johnson NA, Slack GW, Savage KJ, Connors JM, Ben-Neriah S, Rogic S, Scott DW, Tan KL, Steidl C, Sehn LH, Chan WC, Iqbal J, Meyer PN, Lenz G, Wright G, Rimsza LM, Valentino C, Brunhoeber P, Grogan TM, Braziel RM, Cook JR, Tubbs RR, Weisenburger DD, Campo E, Rosenwald A, Ott G, Delabie J, Holcroft C, Jaffe ES, Staudt LM, Gascoyne RD. Concurrent expression of MYC and BCL2 in diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. *J Clin Oncol*. 2012;30(28):3452-3459.
34. Staiger AM, Ziepert M, Horn H, Scott DW, Barth TFE, Bernd HW, Feller AC, Klapper W, Szczepanowski M, Hummel M, Stein H, Lenze D, Hansmann ML, Hartmann S, Moller P, Cogliatti S, Lenz G, Trumper L, Loffler M, Schmitz N, Pfreundschuh M, Rosenwald A, Ott G, German High-Grade Lymphoma Study G. Clinical Impact of the Cell-of-Origin Classification and the MYC/ BCL2 Dual Expresser Status in Diffuse Large B-Cell Lymphoma Treated Within Prospective Clinical Trials of the German High-Grade Non-Hodgkin's Lymphoma Study Group. *J Clin Oncol*. 2017;35(22):2515-2526.

Table S8. CLPA Quality control failures by sample age. Cases are grouped by earliest documented date of biopsy. *Cases <2006 were included in our cohort because they were documented in the pathology record between 2006-2018 but the actual date of biopsy was earlier. Absolute number of cases that passed and failed in each group are shown with (%) of total cases. Fisher exact test between all cases before 2015 (n=384) and all cases 2015 or later (n=194), $p<0.0001$.

Assay Quality Control	Unknown Date	<2006*	2006-2008	2009-2011	2012-2014	≥2015
Pass	29 (100)	4 (80)	84 (76)	121 (92)	123 (90)	190 (98)
Fail	0 (0)	1 (20)	26 (24)	11 (8)	14 (10)	4 (2)

Table S9. Candidate Statistical Learning Models for Classification from the Caret Package (https://topepo.github.io/caret/available-models.html)

Method	Model	Name in Caret	Type	Libraries	Tuning Parameters
Tree	Random Forest	rf	CR	randomForest	mtry
Boosting Tree	eXtreme Gradient Boosting	xgbTree	CR	xgboost, plyr	nrounds, max_depth, eta, gamma, colsample_bytree, min_child_weight, subsample
Boosting Tree	Stochastic Gradient Boosting	gbm	CR	gbm, plyr	n.trees, interaction.depth, shrinkage, n.minobsinnode
Linear DA	High Dimensional Discriminant Analysis	hdda	C	HDclassif	threshold, model
Linear DA	Mixture Discriminant Analysis	mda	C	mda	subclasses
Penalized DA	Penalized Discriminant Analysis	pda	C	pda	lamda
Modified DA	Nearest Shrunken Centroids	pam	C	pamr	threshold
Least Squares/DA	Partial Least Squares-DA	pls	CR	pls	ncomp
Nearest Neighbors	k-Nearest Neighbors	knn	CR	knn	K
Network	Neural Network	nnet	CR	nnet	size, decay
Network	Penalized Multinomial Regression	multinom	C	nnet	decay
Kernal Smoothing	Naive Bayes	nb	C	klaR	fL, usekernel, adjust
SVM	SVM with Polynomial Kernel	svmPoly	CR	kernlab	degree, scale, C
SVM	SVM with Radial Basis Function Kernel	svmRadial	CR	kernlab	sigma, C

*DA=Discriminant Analysis, CR=classification and regression, C=classification, SVM=support vector machine

Table S10. Performance characteristics for CLPA-based calling in the blinded validation cohort of biopsies from previously untreated patients. High probability indicates ≥0.6. PPV=positive predictive value NPV=negative predictive value; *calculated comparing factor to remaining in a one versus all approach; †(specificity+sensitivity)/2; brackets indicate 95% confidence interval.

	WHO Classification (high probability/indeterminate)									
Overall cohort (n=163)	AGG BCL	DLBCL	FL	HL	MCL	MZL	NKTCL	NM	TCL	Total
Accuracy	50%	90%	77%	90%	100%	60%	100%	67%	64%	86% [80%,91%]
Balanced Accuracy*†	75%	92%	88%	93%	100%	79%	100%	82%	81%	
Sensitivity*	50%	90%	77%	90%	100%	60%	100%	67%	64%	
Specificity*	100%	95%	99%	96%	99%	98%	99%	98%	99%	
PPV*	100%	93%	83%	76%	94%	50%	93%	73%	78%	
NPV*	99%	93%	98%	99%	100%	99%	100%	97%	97%	
High-probability cohort (n=135)										
Accuracy	50%	95%	82%	100%	100%	---	100%	100%	67%	94% [89%,97%]
Balanced Accuracy*†	75%	96%	90%	99%	100%	---	100%	100%	83%	
Sensitivity*	50%	97%	82%	100%	100%	---	100%	100%	67%	
Specificity*	100%	96%	98%	98%	100%	---	99%	100%	100%	
PPV*	100%	95%	82%	89%	100%	---	93%	100%	100%	
NPV*	99%	97%	98%	100%	100%	---	100%	100%	98%	

Table S11. Assay performance based on tissue characteristics among samples in the validation cohort. (A) Cases are grouped by earliest documented date of biopsy. *Cases <2006 were included in our cohort because they were documented in the pathology record between 2006-2018 but the actual date of biopsy was earlier. Absolute number of cases that were correctly (Yes) or incorrectly classified (No) in each group are shown with (%) of total cases. Fisher exact test between unknown (n=10), 2006-2008 (n=22), 2009-2011 (n=35), 2012-2014 (n=33) and all cases 2015 or later (n=63), p=0.78. (B) Cases are grouped by biopsy site lymph node or secondary lymphoid tissue (SLT) vs extranodal. Absolute number of cases that were correctly (Yes) or incorrectly classified (No) in each group are shown with (%) of total cases. Fisher exact test between nodal/SLT (n=117) and extranodal (n=37), p=0.79.

(A)

Correct CLPA Classification	**Unknown Date**	**<2006***	**2006-2008**	**2009-2011**	**2012-2014**	**≥2015**
Yes	9 (90)	0 (0)	18 (82)	30 (86)	30 (86)	53 (84)
No	1 (10)	0 (0)	4 (18)	5 (14)	3 (9)	10 (16)

(B)

Correct CLPA Classification	**Nodal/Secondary Lymphoid Tissue**	**Extranodal**
Yes	100 (85%)	31 (84%)
No	17 (15%)	6 (16%)

Table S12. Cases from the validation cohort with CLPA-based calls at probability ≥60% that were misclassified compared to IHC.

Case	INCAN Diagnosis	WHO Diagnosis	WHO Category	CLPA Result	Comment
E0039	Classical Hodgkin Lymphoma	PTCL, NOS	TCL	HL	Lymphohistiocytic infiltrate with atypical cells concerning for peripheral T cell lymphoma; Hodgkin/Reed-Sternberg cells not initially seen, however re-review of IHC markers showed rare CD30+, CD15+ (subset), PAX5+ (variable), MUM1+, EBV ISH negative cells consistent with classic Hodgkin lymphoma
E0290	Lymphoproliferative Disorder	FL, Grade 3A	FL	DLBCL	Morphological range of follicular lymphoma with predominantly grade 3, but also grade 1-2; all areas including those with a predominance of large cells are associated with FDCs; findings are consistent with follicular lymphoma
E0431	High grade pleomorphic lymphoma	ALCL, ALK-negative	TCL	HL	Bands of sclerosis with nodules containing large cells showing crescent/"horse-shoe" nuclei; CD30+, CD15+ (subset), CD45+, CD43+, CD4+ (dim), TIA+, granzyme B+, ALK-negative for all B markers; findings are consistent with ALCL, ALK-
E0482	High grade Diffuse Large Cell Lymphoma	PTCL, NOS	TCL	NKTCL	Predominantly necrotic tissue with focal areas of atypical lymphoid cells, with possible angiodestruction; CD3+, CD5-, CD56-, CD8-, TIA-, granzyme B-, perforin-, CD43+, PDL1+; EBV ISH and LMP1 show few scattered cells; findings show a predominant T cell phenotype, however given extensive necrosis antigenicity may be a concern
E0578	Lymphoproliferative Disorder	DLBCL	DLBCL	FL	Sheets of large pleomorphic B cells with 75% Ki67, c-MYC+, not associated with FDCs; findings consistent with DLBCL
E0639	Diffuse Large B Cell Lymphoma	FL, Grade 1-2	FL	DLBCL	Mixture of small and large cells, predominantly small, however there are areas which show poor fixation which may alter cell size interpretation; BCL6+, BCL2+, CD10-, CD5-, c-MYC (<40%), Ki67 not available; findings compatible with follicular lymphoma grade 1-2
E0695	Lymphoproliferative Disorder	B-LBL	Agg BCL	DLBCL	Monotonous, medium sized B cells with open chromatin; CD45+ (dim), CD20 (negative to dim), PAX5+, CD10+, TdT+; findings consistent with B lymphoblastic lymphoma/leukemia
E0913	Diffuse Large B Cell Lymphoma	DLBCL	DLBCL	FL	Sheets of atypical medium sized B cells with 50% Ki67 and no associated FDC networks; findings compatible with DLBCL

Table S13. Full performance characteristics for CLPA-based calling in the cohort of samples from patients with relapsed/refractory lymphoma (n=39). WHO classification for each lymphoma based on H&E, IHC, and when indicated FISH is shown in columns and CLPA calls are in rows. Abbreviations: PPV=positive predictive value, NPV=negative predictive value. *calculated comparing factor to remaining in a one versus all approach; †(specificity+sensitivity)/2; brackets [ ] indicate 95% confidence interval.

	WHO Classification high probability/indeterminate							
CLPA-Predicted	AGG BCL	DLBCL	FL	HL	MCL	MZL	NK-TCL	Total
Agg BCL	0/0							0
DLBCL	1/0	17/0	2/0			1/0		21
FL		0/1	2/1					4
HL				5/0				5
MCL			0/1		5/0			6
MZL						0/0		0
NK-TCL							1/0	1
Nonmalignant		0/1						1
TCL				0/1				1
Total Overall	1	19	6	6	5	1	1	39
Accuracy	0%	89%	50%	83%	100%	0%	100%	79% [64%,91%]
Balanced Accuracy*†	50%	85%	73%	92%	99%	50%	100%	
Sensitivity*	0%	89%	50%	83%	100%	0%	100%	
Specificity*	100%	80%	97%	100%	97%	100%	100%	
PPV*	---	81%	75%	100%	83%	---	100%	
NPV*	97%	89%	91%	97%	100%	97%	100%	
Total indeterminate	0 (0%)	2 (11%)	2 (33%)	1 (17%)	0 (0%)	0 (0%)	0 (0%)	5 (13%)
High Probability cohort								
Total Evaluable	1	17	4	5	5	1	1	34
Accuracy	0%	100%	50%	100%	100%	0%	100%	88% [73%,97%]
Balanced Accuracy*†	50%	88%	75%	100%	100%	50%	100%	
Sensitivity*	0%	100%	50%	100%	100%	0%	100%	
Specificity*	100%	76%	100%	100%	100%	100%	100%	
PPV*	---	81%	100%	100%	100%	---	100%	
NPV*	97%	100%	94%	100%	100%	97%	100%	

Table S14. Detailed cost analysis for assay reagents and supplies in Guatemala. A. Analysis conditions **B.** Quantity of reagents per reactions **C**. Quantity of reagents per 96 reactions **D.** Laboratory consumables **E.** Additional reagents **F.** Total assay cost. Prices are listed in Guatemalan quetzale (Q) and U.S. dollars ($)

A

Runs of 95 samples + 1 negative control / week
Using the cartridge version of 6 months lifetime
Using 3 Aluminum foil per run (96 reactions)
Using 2 Plastic foil per run (96 reactions)
Using 1 cathode buffer container every two weeks, because that is the cathode buffer lifetime
Using 16 Septa for cathode container buffer every two weeks
Using 96 Septa for plate every run
Using 10 uL tips to introduce the scrolls
The price in dollars was calculated with the official exchange rate of the Bank of Guatemala at 12/10/2019

B

Reagent	uL/well
Lymphoma Mix B	15
Lymphoma Mix A	15
DirectMix C	5
DxBuffer1	50
Buffer TE	17
Direct Beads	5
Direct Wash	540
DirectTaq	5
DxPrime	5
GeneScan 600LIZ	0.5
Formamide	17.5

C

Reagent	uL /96 reactions
Lymphoma Mix B	1444.5
Lymphoma Mix A	1444.5
DirectMix C	481.5

DxBuffer1	4815
Buffer TE	1637.1
Direct Beads	481.5
Direct Wash	52000
DirectTaq	481.5
DxPrime	481.5
GeneScan 600LIZ	48.15
Formamide	1685.25

D

Lab Consumable	Units of measurement (UOM)	UOM Qty	Price per UOM (Local currency)	96 Units/reaction	Price per 96 reactions (Local currency)	Price per 96 reactions ($)
96 well PCR plate	Pack	10	Q 663.08	2	Q 132.62	$16.81
2 mL tubes	Pack	250	Q 205.00	193	Q 158.26	$20.06
Cathode Buffer Container (125 injections per container)	Box	4	Q 1,413.96	0.5	Q 176.75	$22.40
Seq Plate	Box	10	Q 663.08	1	Q 66.31	$8.40
10 uL tips	pack	960	Q 731.28	480	Q 365.64	$46.34
20 uL tips	pack	960	Q 731.28	384	Q 292.51	$37.07
100 uL tips	pack	960	Q 731.28	288	Q 219.38	$27.81
200 uL tips	pack	960	Q 731.28	864	Q 658.15	$83.42
1000 uL tips	pack	960	Q 731.28	6	Q 4.57	$0.58
Cartridge (1000 reactions)	box	1000	Q 19,219.20	96	Q 1,845.04	$233.85
Reagent reservoir	pack	10	Q 186.00	1	Q 18.60	$2.36
Plastic Foil	box	100	Q 1,553.88	2	Q 31.08	$3.94
16 Septa for Cathode buffer container	pack	20	Q 2,814.48	1	Q 140.72	$17.84
Septa for Plate	pack	20	Q 3,700.00	1	Q 185.00	$23.45
Aluminum Foil	pack	100	Q 2,016.00	3	Q 60.48	$7.67
Gloves	box	100	Q 55.00	8	Q 4.40	$0.56
Consumable price per 96 reactions					**Q 4,359.51**	**$552.54**

E

Reagent	Units of measurement (UOM)	UOM Qty in uL	Price per UOM (Local currency)	uL used per 96 reactions	Price per 96 reactions (Local currency)	Price per 96 reactions ($)
GeneScan 600LIZ	2 tubes of 200 uL each	400	Q 4,526.11	48.15	Q 544.83	$69.05
Formamide	bottle	5000	Q 471.74	1685.25	Q 159.00	$20.15

Reagent price per 16 reactions					**Q 703.83**	$89.21

F

	Local Currency	U.S. Dollars
Total Cost per 96 reactions	Q 5,063.34	$641.74
Cost per patient (in runs of 95 patients + 1 negative control)	**Q 53.30**	**$6.76**

Supplementary Figure Legends

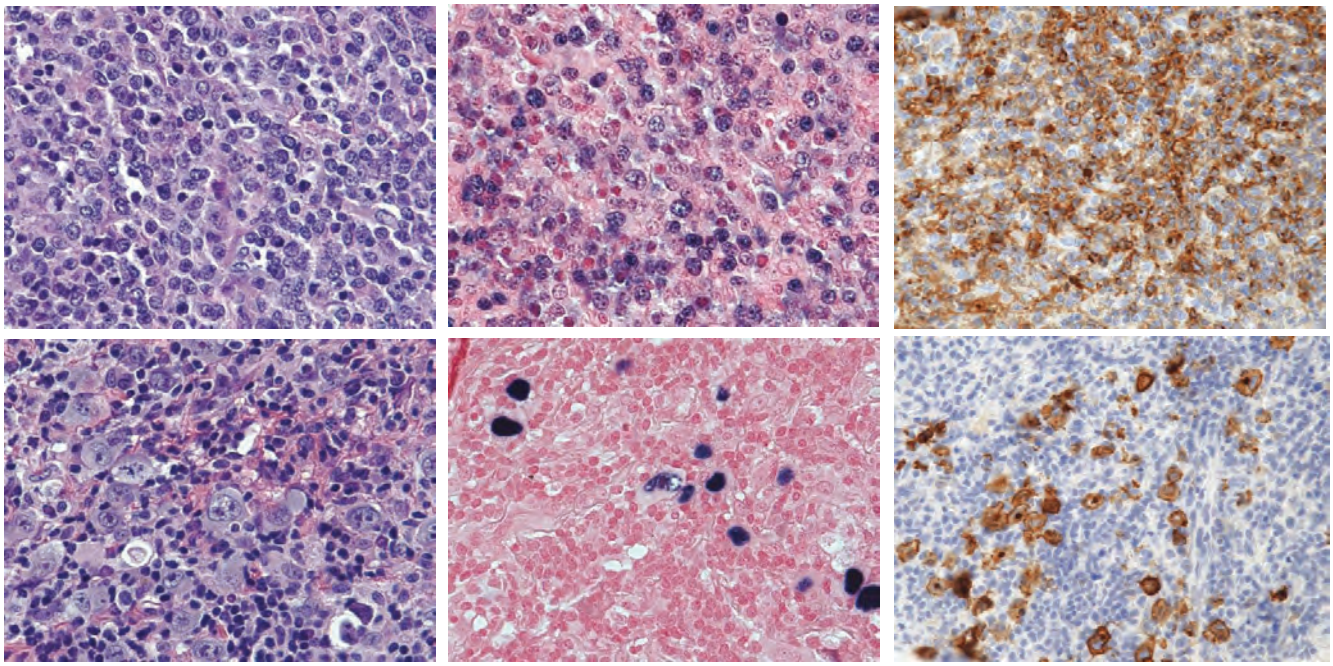
Figure S1. IHC and gene expression assessment. **A.** Examples of assessments from tissue microarray (TMA). Top left, H&E of EBV+ Diffuse large B cell lymphoma (DLBCL). Top middle, *in situ* hybridization for EBV in DLBCL. Top right, CD20 expression in EBV+ DLBCL. Bottom left, H&E of classic Hodgkin Lymphoma (cHL). Bottom middle, *in situ* hybridization for EBV in cHL. Bottom right, CD30 expression in cHL. **B.** CLPA assay measures gene expression directly from 6-10 um FFPE section through same tube RNA isolation and by chemical ligation of probes bound to the RNA followed by streptavidin-conjugated paramagnetic bead-capture. PCR is performed so that each product has a unique size. Quantification of the size-differentiated, ligated probe amplification products is performed using a capillary electrophoresis instrument. SS, spacer sequence; TS, target sequence; UFP, universal forward primer; URP, universal reverse primer

Figure S2. A. Histograms of normalized CLPA expression for each gene for all pre-treatment cases in the test and validation sets (n=560) **B.** Boxplots of normalized CLPA expression values by WHO diagnosis. Horizontal lines indicate the median, boxes represent the interquartile range (IQR) with error bars extending to 1.5 times the IQR, and dots indicating value >1.5 time the IQR. **C.** Spearman correlation plot matrix where blue indicates positive correlation coefficient and red indicates negative correlation coefficient. Color intensity and the size of the circle are proportional to the correlation coefficients.

Figure S3. A. Spearman correlation plot matrix for the relapsed cases (n=39) where blue indicates positive correlation coefficient and red indicates negative correlation coefficient. Color intensity and the size of the circle are proportional to the correlation coefficients.

B. Boxplots of normalized CLPA expression values by diagnosis for the relapse cases (n=39). Horizontal lines indicate the median, boxes represent the interquartile range (IQR) with error bars extending to 1.5 times the IQR, and dots indicating value >1.5 time the IQR.

A



B

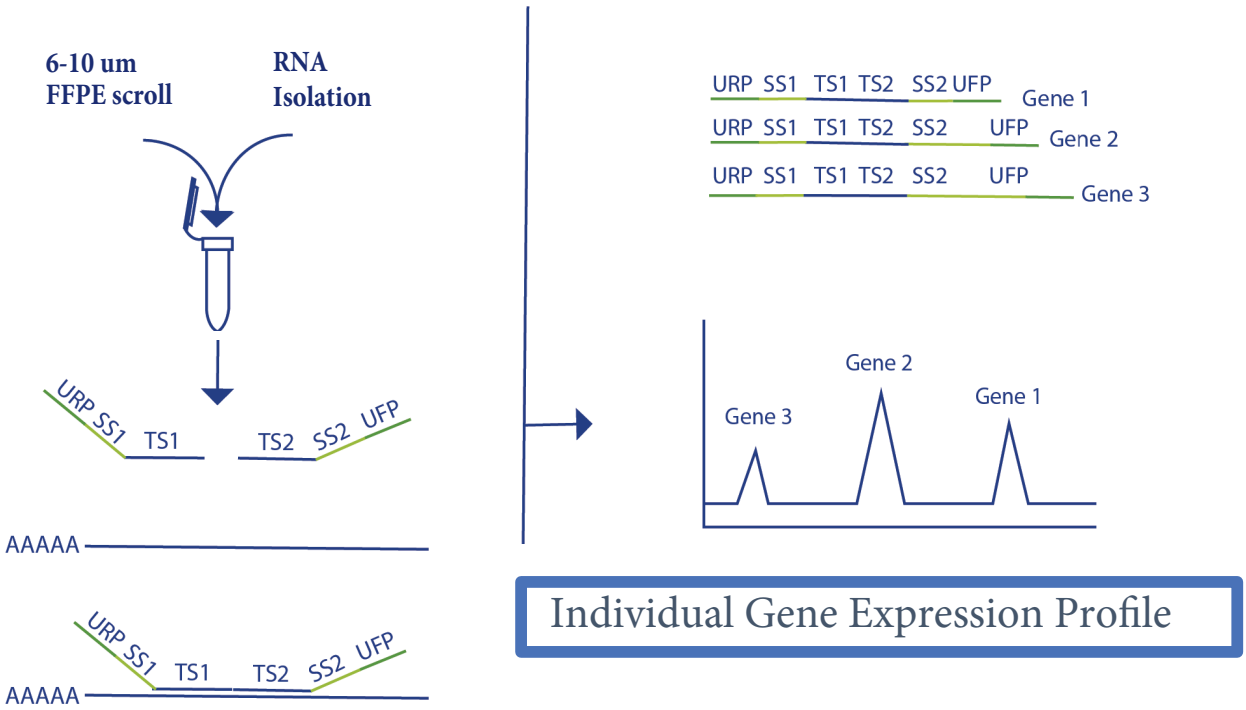

6-10 um
FFPE scroll
RNA
Isolation
URP SS1 TS1 TS2 SS2 UFP
AAAAA
URP SS1 TS1 TS2 SS2 UFP
AAAAA
URP SS1 TS1 TS2 SS2 UFP Gene 1
URP SS1 TS1 TS2 SS2 UFP Gene 2
URP SS1 TS1 TS2 SS2 UFP Gene 3
Gene 3
Gene 2
Gene 1
Individual Gene Expression Profile

Figure S1

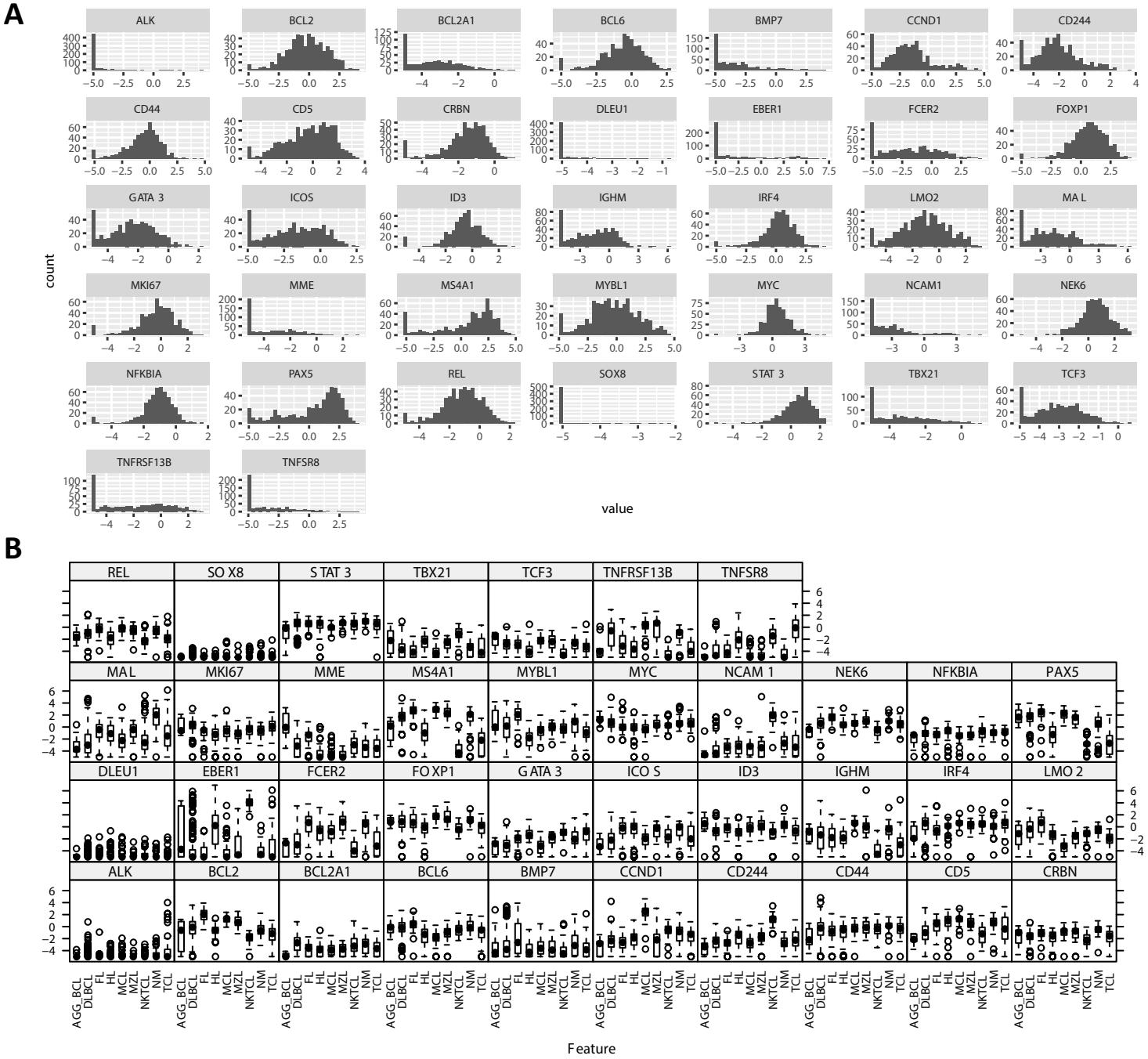
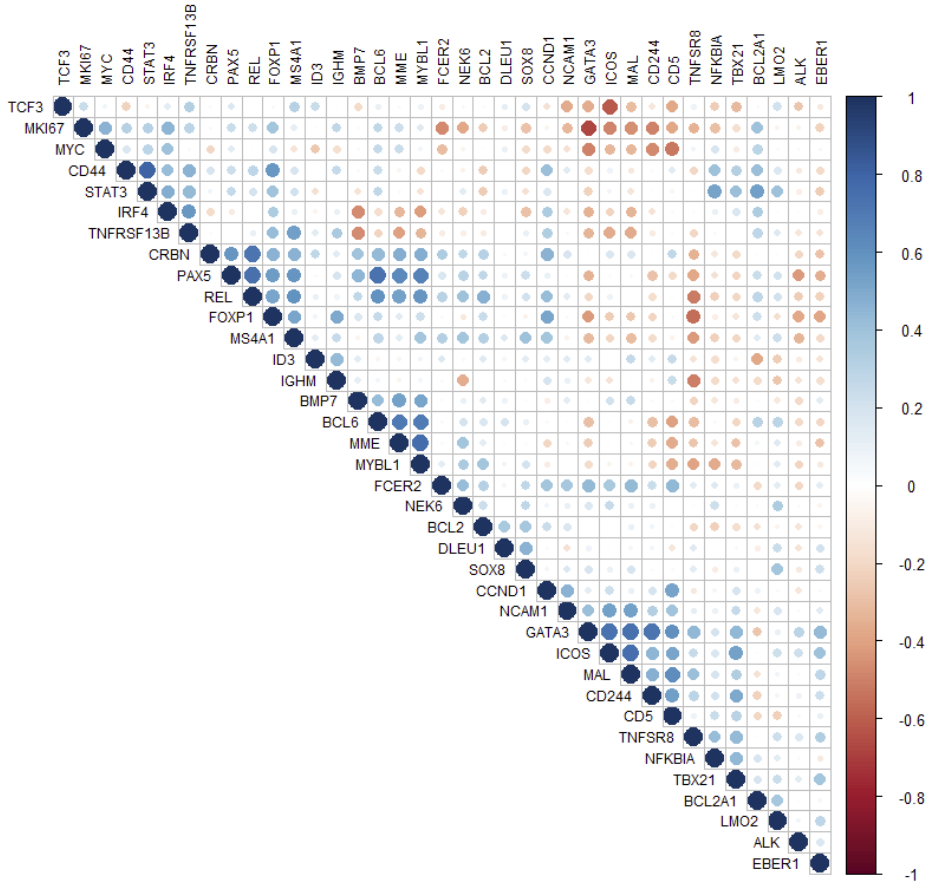

A
ALK
BCL2
BCL2A1
BCL6
BMP7
CCND1
CD244
CD44
CD5
CRBN
DLEU1
EBER1
FCER2
FOXP1
GATA 3
ICOS
ID3
IGHM
IRF4
LMO2
MA L
MKI67
MME
MS4A1
MYBL1
MYC
NCAM1
NEK6
NFKBIA
PAX5
REL
SOX8
STAT 3
TBX21
TCF3
TNFRSF13B
TNFSR8
count
value
B
REL
SO X8
S TAT 3
TBX21
TCF3
TNFRSF13B
TNFSR8
MA L
MKI67
MME
MS4A1
MYBL1
MYC
NCAM 1
NEK6
NFKBIA
PAX5
DLEU1
EBER1
FCER2
FO XP1
G ATA 3
ICO S
ID3
IGHM
IRF4
LMO 2
ALK
BCL2
BCL2A1
BCL6
BMP7
CCND1
CD244
CD44
CD5
CRBN
AGG_BCL
DLBCL
FL
HL
MCL
MZL
NKTCL
NM
TCL
Feature
C
NFKBIA
GATA3
TNFSR8
DLEU1
SOX8
ALK
EBER1
NCAM1
CD244
TBX21
MKI67
MYC
IRF4
IGHM
FOXP1
TNFRSF13B
FCER2
CD5
ICOS
MAL
BCL2A1
CD44
STAT3
CCND1
CRBN
ID3
TCF3
BCL2
REL
MS4A1
PAX5
BMP7
LMO2
NEK6
MYBL1
BCL6
MME

Figure S2

A


TCF3
MKI67
MYC
CD44
STAT3
IRF4
TNFRSF13B
CRBN
PAX5
REL
FOXP1
MS4A1
ID3
IGHM
BMP7
BCL6
MME
MYBL1
FCER2
NEK6
BCL2
DLEU1
SOX8
CCND1
NCAM1
GATA3
ICOS
MAL
CD244
CD5
TNFSR8
NFKBIA
TBX21
BCL2A1
LMO2
ALK
EBER1
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

B

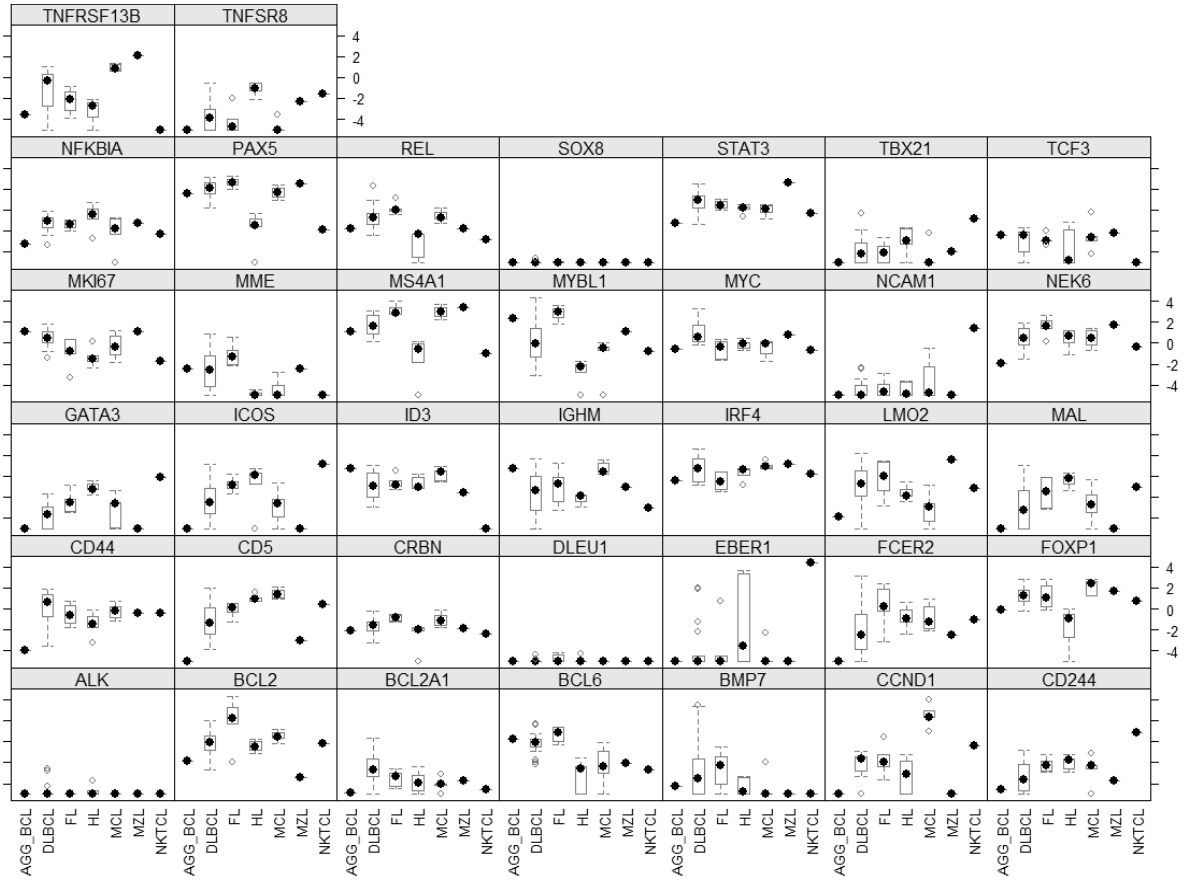

TNFRSF13B
TNFSR8
NFKBIA
PAX5
REL
SOX8
STAT3
TBX21
TCF3
MKI67
MME
MS4A1
MYBL1
MYC
NCAM1
NEK6
GATA3
ICOS
ID3
IGHM
IRF4
LMO2
MAL
CD44
CD5
CRBN
DLEU1
EBER1
FCER2
FOXP1
ALK
BCL2
BCL2A1
BCL6
BMP7
CCND1
CD244
4
2
0
-2
-4
AGG_BCL
DLBCL
FL
HL
MCL
MZL
NKTCL

Figure S3