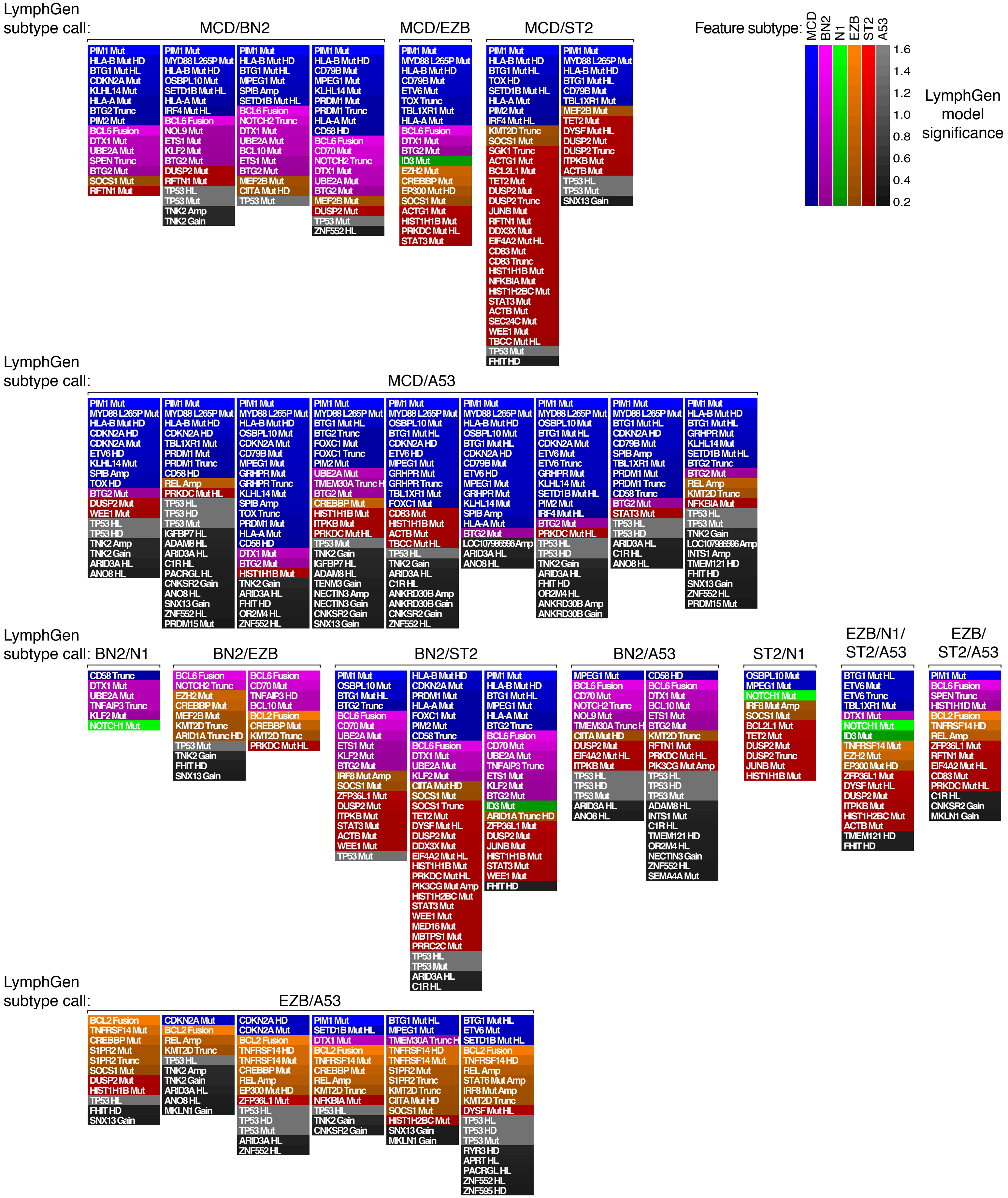

Figure S1 (related to Figure 1) Genetic features in genetically composite DLBCL tumors. Shown are genetic features in DLBCL tumors that have been classified into more than one subtype by the LymphGen algorithm (n=33). The significance of each genetic feature in the LymphGen model is presented according to the color scale at the right.

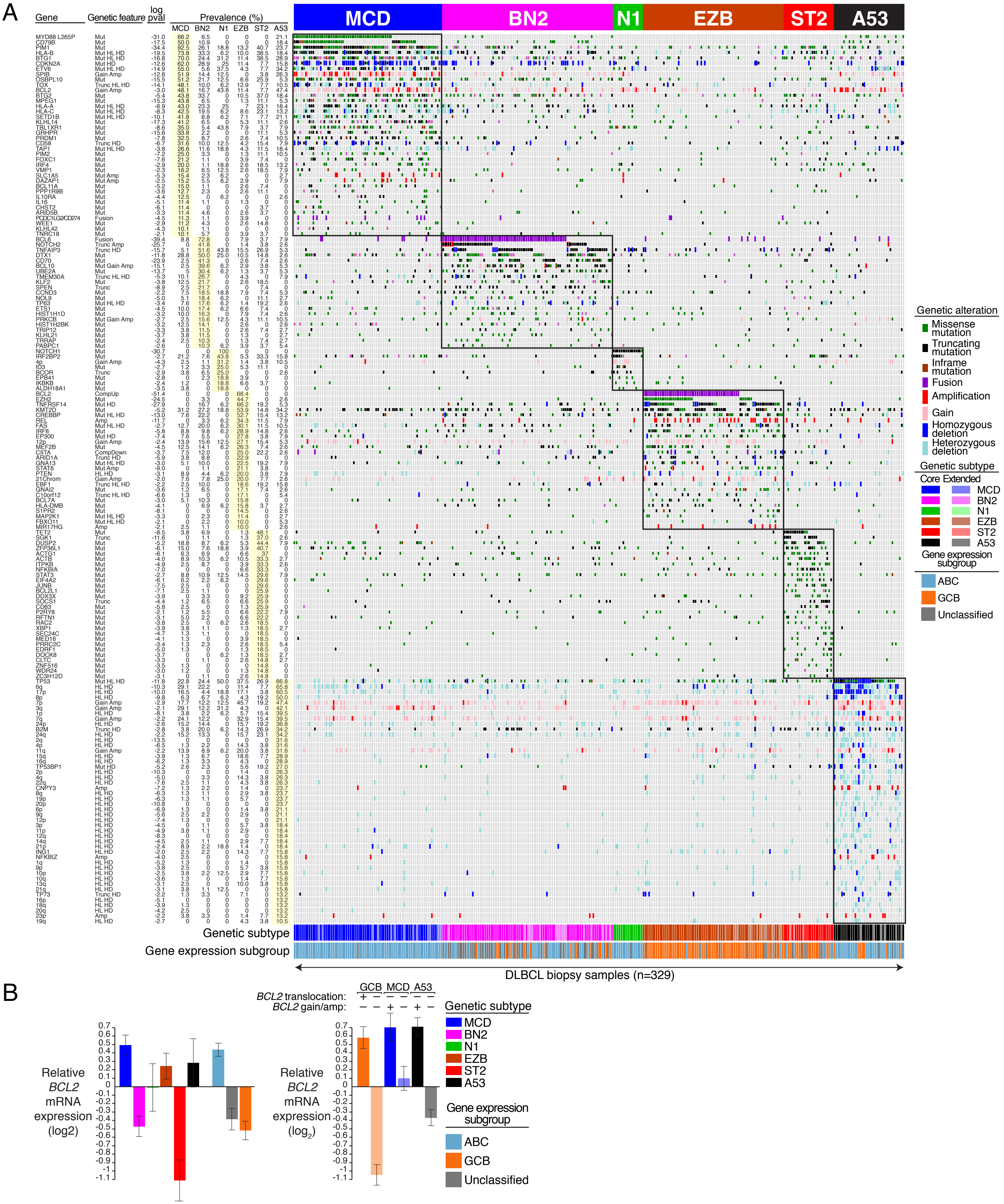

Figure S2 (related to Figure 2). A. Genetic features that distinguish DLBCL subtypes. Shown are genetic features in DLBCL biopsy samples that have been classified into a single subtype (n=329) according to the legend at the right. Also shown at bottom are the subtype calls (core vs. extended) and the COO subgroup associated with each case, according to the legend at the right. **B.** Expression of *BCL2* mRNA in the indicated subsets of DLBCL. Each value is the difference between the average *BCL2* expression in the subset of cases relative to the mean expression over all cases. Left panel displays expression in genetic subtypes and gene expression subgroups. Right panel displays the influence of *BCL2* translocation or *BCL2* gain/amplification on expression levels in the indicated subset of cases.

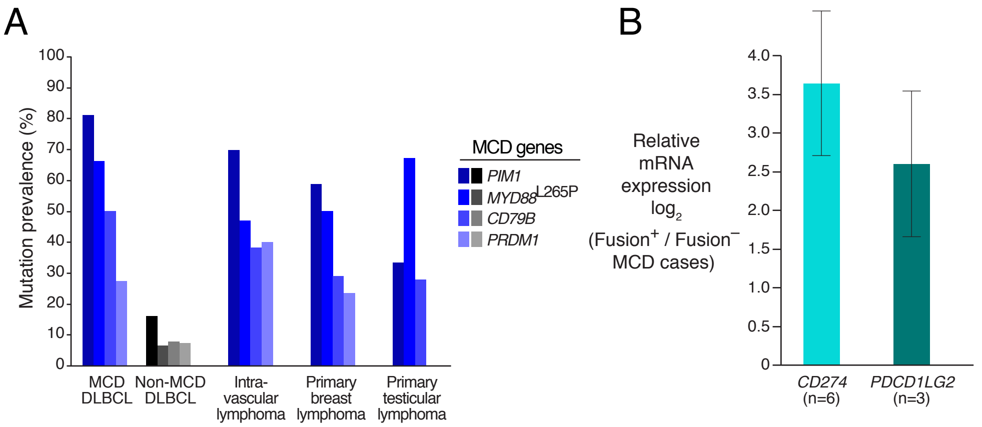

Figure S3 (related to Figure 3). A. Prevalence of the indicated MCD-defining mutations in primary testicular, primary breast and primary intravascular lymphomas. **B.** Relative expression of *CD274* mRNA, encoding PD-L1, and *PDCD1LG2* mRNA, encoding PD-L2, in MCD tumors with fusions of the respective genes versus other, fusion-negative MCD tumors.

A

	NCI Cohort (Schmitz et al)	Harvard Cohort (Chapuy et al)	BCC Cohort (Ennishi et al)
# Patients	574	304	332
# Patients with survival	240	263	332
FFPE cases	0%	45%	0%
Exome	+	+	−
Panel seq.	359 genes	−	82 genes
CGH	Whole genome	Regional	Whole genome
Translocation	+	+	+
COO	+	+	+

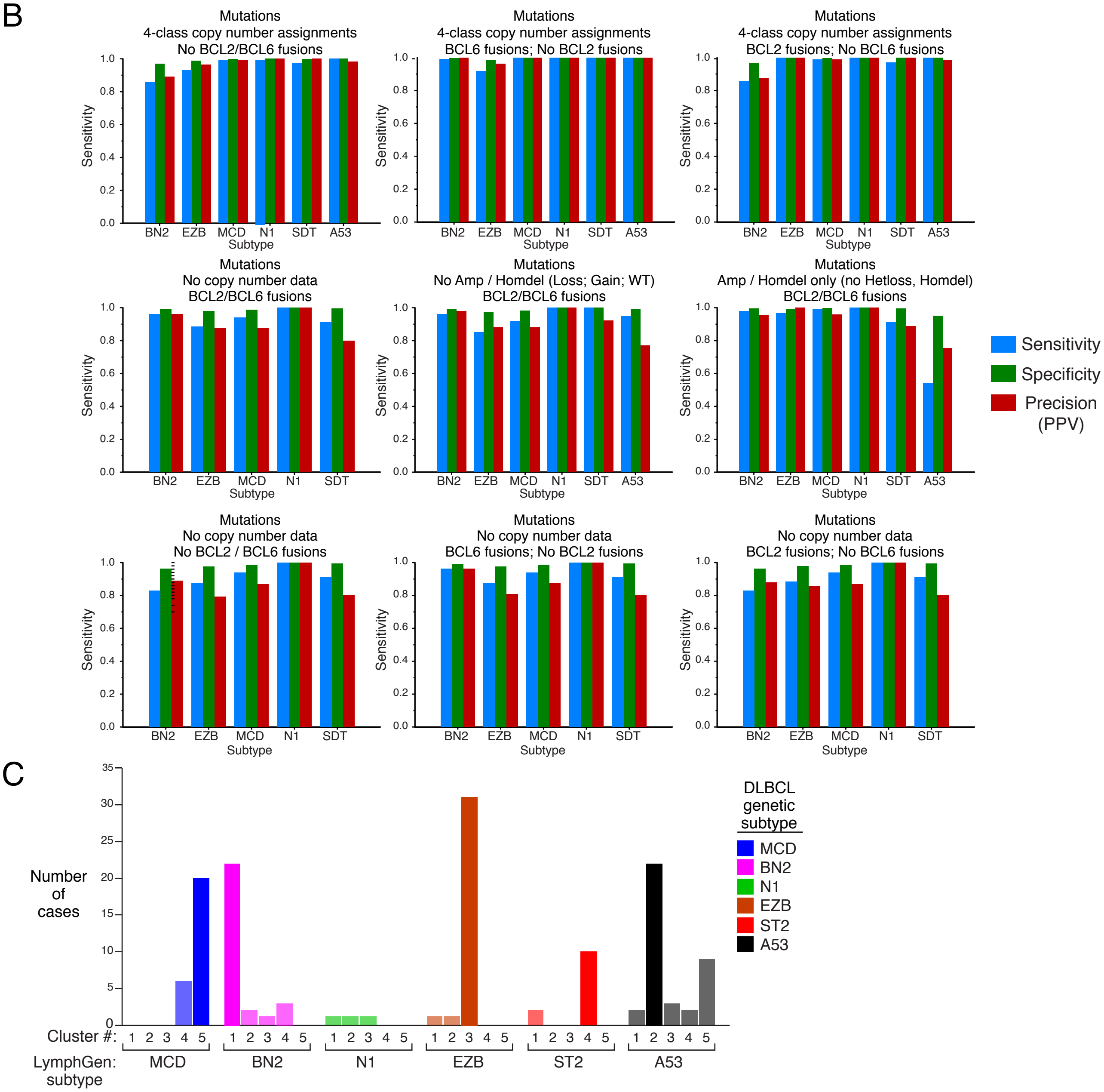

Figure S4 (related to Figure 4). A. Clinical and analytic aspects of the three DLBCL cohorts used in the present study (Chapuy et al., 2018; Ennishi et al., 2019a; Schmitz et al., 2018). **B.** Sensitivity, specificity and precision (positive predictive value (PPV)) for LymphGen predictor models based on the indicated availability of dataset types. Each plot is based on data from the NCI cohort and is a comparison with the LymphGen model based on a full complement of genetic data (mutations, 4-class copy number assignments (amplification, gain, heterozygous loss, homozygous deletion), and fusions (BCL2, BCL6). Models lacking copy number data did not include A53 since this subtype is defined primarily by copy number alterations. **C.** Relationship between clusters defined in the Harvard cohort (Chapuy et al., 2018) and LymphGen-predicted subtypes. All LymphGen subtypes were predominantly composed of cases belonging to a single cluster, with the exception of cases in the N1 subtype, which were not assigned to a cluster.

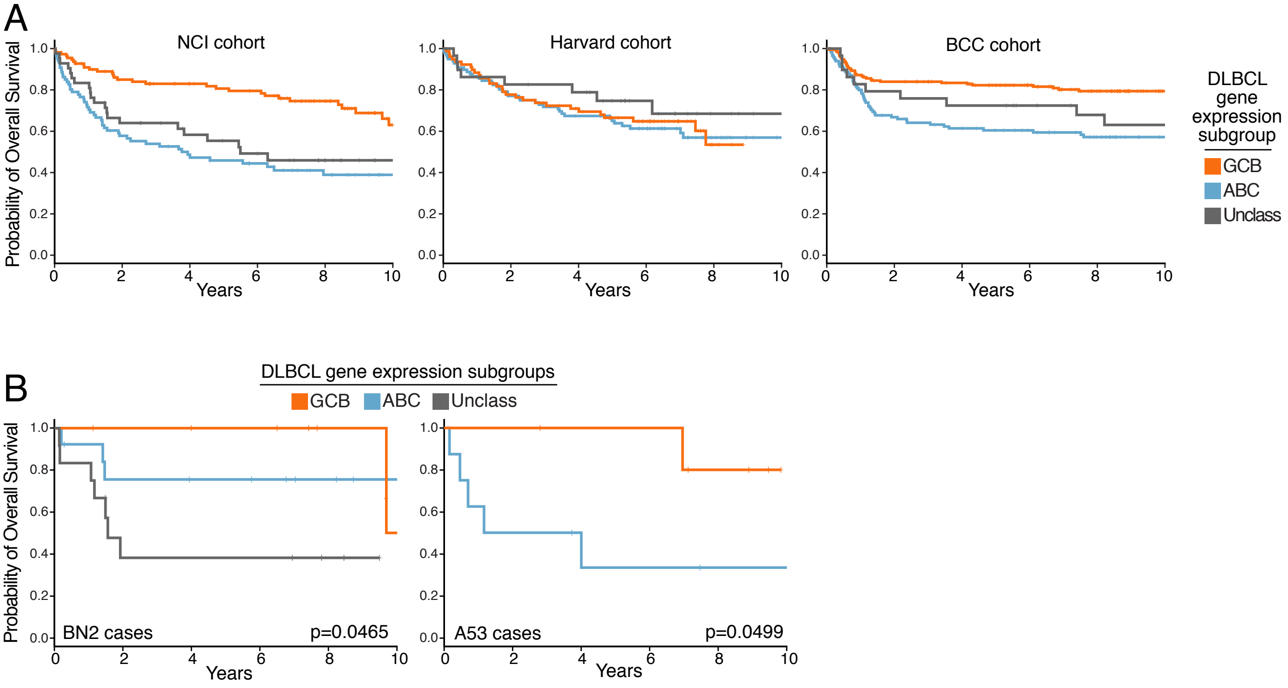

Figure S5 (related to Figure 4). A. Kaplan-Meier plots of overall survival in the indicated DLBCL cohorts according to COO gene expression subgroup. **B.** Kaplan-Meier plots of overall survival in the BN2 and A53 genetic subtypes according to COO gene expression subgroup within the NCI cohort.

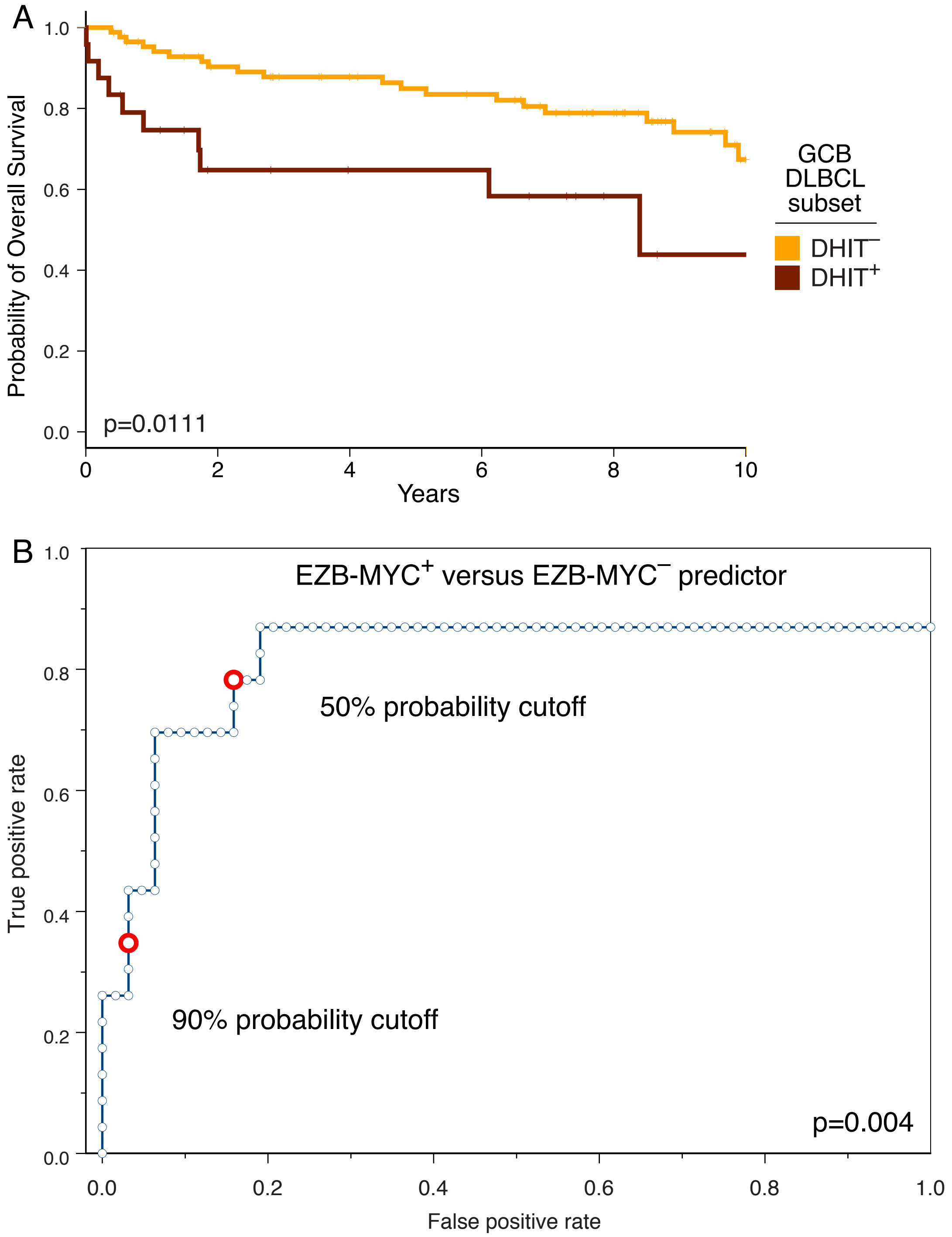

Figure S6 (related to Figure 6). A. Kaplan-Meier plot of overall survival of DHIT+ and DHIT– subsets of GCB DLBCL. **B.** Receiver operating characteristic (ROC) curve for the binary probabilistic classifier of EZB-M^+ versus EZB-M^-.

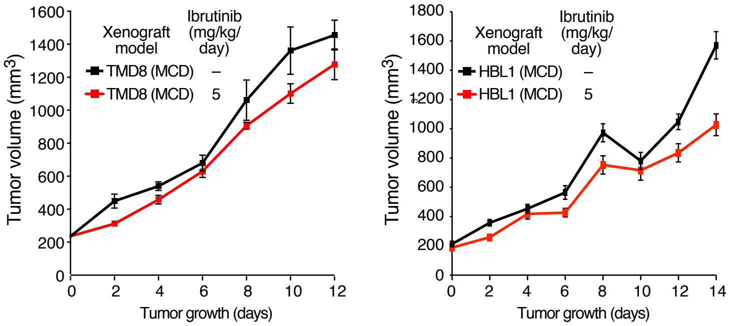

Figure S7 (related to Figure 7). Effect of low doses of ibrutinib (5 mg/kg/day) on growth of xenografts of two MCD models, TMD8 (left), and HBL1 (right).

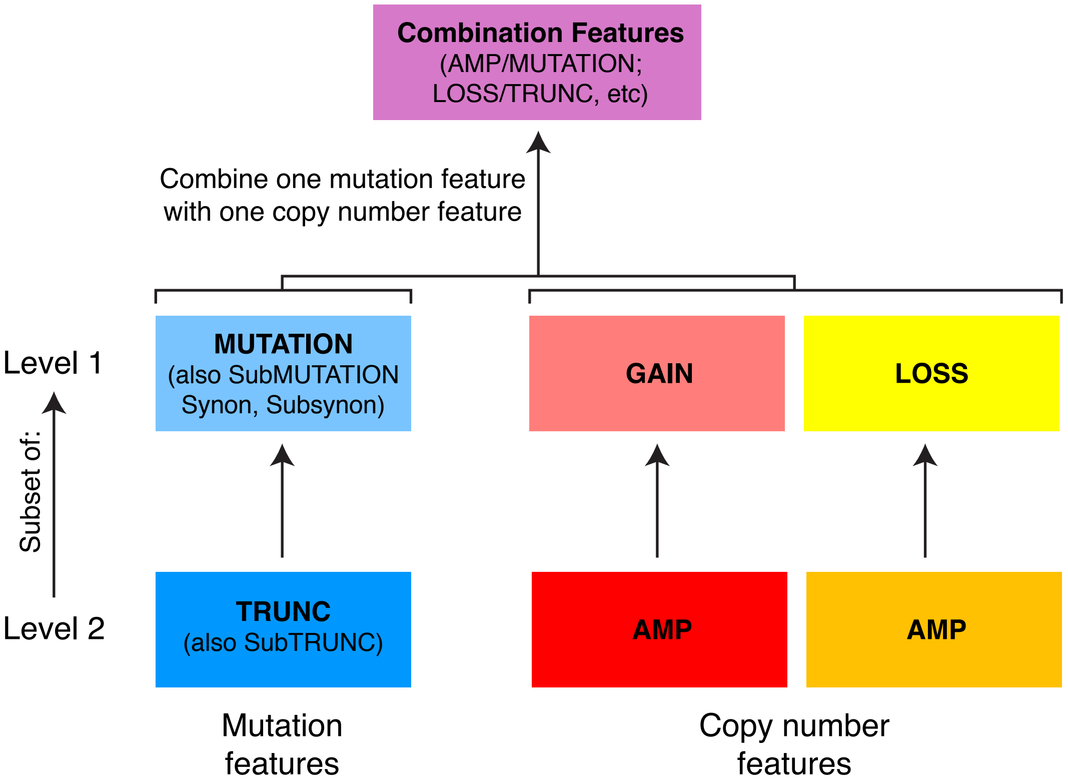

Figure S8 (related to STAR methods). Hierarchical organization of genetic features for use in the LymphGen algorithm.!

Table S3: Statistical analysis of subtype-defining genetic features in the Harvard and BCC cohorts. Related to Figure 4

Probability that subtype-defining genetic features co-occur

Subtype-defining features evaluated	**Harvard cohort**	**BCA cohort**
BN2	5.97E-19	1.28E-09
EZB	8.03E-16	3.21E-12
MCD	1.13E-30	2.24E-19
SDT	6.34E-07	5.08E-13

Probability that TP53 mutation and/or deletion is associated with more copy number abnormalities

Harvard cohort	**BCC cohort**
1.07E-17	4.37E-09

Table S5: Gene expression signature averages and statistical associations with genetic subtypes. Related to Figure 6

			Mean Value of Signature average within Subtype							P-value for difference in signature average between those within vs outside subtype						
Short signature name	**Signature**	**Signature category**	**MCD**	**BN2**	**N1**	**EZB-MYC-**	**EZB-MYC+**	**SDT**	**A53**	**MCD**	**BN2**	**N1**	**EZB-MYC-**	**EZB-MYC+**	**SDT**	**A53**
Prolif-11	MCL_proliferation_survival	Malignant processes	0.349	0.009	-0.656	0.492	-0.475	0.009	-0.173	2.80E-06	6.62E-02	8.69E-02	8.05E-03	6.22E-02	3.63E-01	9.41E-01
MycUp-4	Myc_ChIP_PET_Expr_Up	Malignant processes	0.434	0.037	-0.522	0.738	-0.536	-0.074	0.008	4.29E-07	1.66E-01	9.70E-02	4.67E-04	1.74E-03	8.99E-01	5.06E-01
Ribo-1	Ribosomal_protein	Malignant processes	0.313	0.044	0.260	1.010	0.156	0.200	0.122	1.20E-01	2.68E-01	6.59E-01	8.42E-05	7.56E-01	7.97E-01	8.59E-01
Quiesce-2	Quiescence_heme_cluster1	Malignant processes	-0.209	0.018	0.755	-1.161	-0.041	-0.135	-0.280	2.75E-02	9.81E-01	2.98E-03	7.44E-08	3.89E-01	4.23E-01	5.69E-02
Glycol-1	Glycolysis	Malignant processes	-0.045	0.155	-0.407	0.525	0.116	0.600	-0.212	7.67E-01	9.17E-02	1.29E-01	1.52E-02	2.58E-01	2.25E-03	2.27E-01
Lipid-1	Lipid_enzyme_Node3152	Malignant processes	-0.310	-0.002	-0.292	0.916	0.263	-0.019	-0.052	3.05E-04	6.32E-01	1.54E-01	1.06E-05	2.92E-02	7.39E-01	5.29E-01
GCB-1	GC_B_cell_U133Plus	B cell differentiation	-0.414	0.414	-0.554	1.407	1.200	0.604	-0.150	7.59E-10	4.40E-02	1.90E-03	2.83E-10	3.93E-17	4.81E-02	1.96E-02
GCB-8	scGC_cluster1_LZ	B cell differentiation	-0.211	0.122	-0.249	0.185	0.566	0.356	-0.034	6.28E-02	3.02E-01	3.61E-01	3.20E-01	9.60E-05	1.14E-01	8.09E-01
GCB-9	scGC_cluster1_IZ	B cell differentiation	0.149	0.111	0.072	1.386	0.415	0.162	0.022	8.92E-01	8.18E-01	8.12E-01	3.99E-09	1.09E-02	8.88E-01	4.96E-01
GCB-10	scGC_cluster1_DZ	B cell differentiation	0.319	0.178	-0.194	1.146	0.085	0.152	-0.009	3.48E-02	3.78E-01	2.73E-01	2.05E-06	7.89E-01	7.60E-01	5.49E-01
PC-1	Plasma_cell_gt_mature_Bcell	B cell differentiation	-0.498	0.077	0.264	-1.247	-0.035	0.385	-0.340	3.46E-05	1.13E-01	1.71E-01	8.35E-08	9.58E-01	1.67E-02	8.88E-02
Bcell-6	Memory_B_CCR6pos_Memory_B_precursor_gt_DZ_CCR6neg_LZ	B cell differentiation	-0.023	0.040	1.142	-1.415	-0.697	-0.443	-0.076	5.73E-01	2.14E-01	1.80E-06	2.33E-10	1.59E-07	7.52E-02	9.60E-01
IRF4Up-7	IRF4_ABC_induced_ABCDLBCL	B cell transcription factors	0.707	0.041	0.100	-2.145	-1.070	-0.729	0.278	3.60E-18	6.13E-02	3.33E-01	9.81E-28	4.96E-21	1.16E-03	6.70E-03
IRF4Dn-1	IRF4_ABC_repressed_GCBDLBCL	B cell transcription factors	-0.658	0.365	-0.519	1.472	1.161	0.627	-0.204	5.55E-15	2.83E-02	8.18E-03	5.46E-12	4.02E-18	1.56E-02	2.76E-02
OCT2Up-1	OCT2_shRNA_OCT2_ChIP	B cell transcription factors	0.403	0.121	-0.776	0.517	-0.553	-0.297	-0.089	1.72E-07	3.95E-03	2.47E-02	5.73E-03	1.08E-02	5.60E-01	6.11E-01
BCL6Dn-1	BCL6_repressed	B cell transcription factors	0.300	-0.032	0.502	-1.367	-0.317	-0.181	-0.403	1.57E-03	9.44E-01	3.50E-02	1.99E-09	9.24E-03	4.77E-01	2.55E-02
TCF3Up-1	TCF3_BL_up	B cell transcription factors	-0.552	0.111	-0.473	0.368	0.421	0.322	-0.378	2.17E-08	2.44E-01	4.73E-02	4.94E-02	4.59E-04	8.71E-02	1.24E-02
NFkB-10	NFkB_Up_HBL1	Oncogenic signaling pathways	0.332	0.169	-0.220	-2.612	-0.315	0.024	-0.114	1.24E-04	8.21E-03	7.12E-01	2.74E-26	2.14E-02	5.20E-01	9.92E-01
p53up-1	p53_up_Xray	Oncogenic signaling pathways	0.123	0.394	0.141	0.459	0.266	0.368	-0.706	6.36E-01	1.20E-03	7.90E-01	7.15E-02	1.14E-01	1.40E-01	1.77E-06
NotchUp-6	Notch_CLL_up_2x_Fabbri	Oncogenic signaling pathways	-0.651	0.203	0.721	-0.984	0.279	0.211	-0.393	3.06E-10	7.33E-03	1.19E-03	1.25E-05	1.73E-02	1.71E-01	1.82E-02
PI3KUp-1	PI3K_overexpression_up	Oncogenic signaling pathways	0.249	0.162	-0.079	-0.044	0.401	0.823	-0.289	2.17E-01	5.39E-01	5.89E-01	7.26E-01	5.39E-02	2.16E-03	5.47E-02
JAK2Up-2	JAK2_PMBL_gt_GCB	Oncogenic signaling pathways	-0.128	0.114	0.336	-1.504	0.189	0.796	-0.544	2.57E-01	3.26E-01	2.44E-01	1.39E-09	3.74E-01	3.12E-04	2.12E-03
JAKUp-1	JAK_IL10_Ly10_Up	Oncogenic signaling pathways	0.319	0.079	-0.052	-1.978	-0.355	-0.353	-0.268	1.18E-05	1.96E-02	6.92E-01	5.55E-16	3.06E-02	3.27E-01	4.95E-01
GCThUp-1	GC_T_helper_up_Chtanova_and_Kim	Tumor microenvironment	-0.534	0.123	0.558	0.199	0.593	0.032	-0.586	6.44E-07	2.57E-01	3.60E-02	2.66E-01	1.07E-05	9.06E-01	3.43E-04
CD4T-2	CD4_T_cell_naive_Newman	Tumor microenvironment	-0.285	-0.050	0.534	-0.662	0.156	-0.072	-0.508	4.61E-03	6.75E-01	1.74E-02	2.25E-03	2.20E-01	7.41E-01	6.48E-04
CD8T-7	CD8_T_cell_Newman	Tumor microenvironment	-0.186	-0.186	0.373	-0.704	0.032	-0.178	-0.560	2.82E-01	2.42E-01	6.17E-02	5.07E-03	4.89E-01	5.90E-01	1.53E-03
Treg-1	Regulatory_T_cell_McHugh_Herman_concensus	Tumor microenvironment	0.050	-0.107	0.281	-1.505	0.053	0.097	-0.581	3.00E-01	7.26E-01	2.03E-01	8.98E-09	6.18E-01	4.36E-01	3.44E-03
NK-3	NK_cell_resting_Newman	Tumor microenvironment	-0.094	-0.176	0.454	-0.953	0.019	-0.262	-0.501	7.72E-01	2.39E-01	3.58E-02	9.27E-05	6.96E-01	3.02E-01	5.23E-03
MPhage-1	Macrophage_M1_Newman	Tumor microenvironment	0.159	0.157	0.347	-0.874	0.053	0.204	-0.346	3.48E-01	3.21E-01	2.46E-01	2.09E-05	7.12E-01	4.60E-01	8.46E-03
MPhage-2	Macrophage_M2_Newman	Tumor microenvironment	-0.013	0.098	0.357	-0.853	0.009	0.083	-0.463	9.22E-01	2.80E-01	1.38E-01	9.02E-05	8.65E-01	6.48E-01	2.75E-03
DC-9	Dendritic_cell_activated_Newman	Tumor microenvironment	-0.241	0.105	0.149	-0.979	0.089	0.207	-0.509	7.70E-02	8.19E-02	3.97E-01	4.17E-05	3.62E-01	1.62E-01	3.76E-03
Stromal-1	Stromal-1_DLBCL_survival_predictor	Tumor microenvironment	-0.522	0.195	-0.083	-0.126	0.705	0.414	-0.289	1.09E-07	1.12E-01	6.36E-01	6.93E-01	3.19E-07	5.80E-02	4.28E-02

Table S6: Subtype-defining genetic features in DLBCL subtype cell line models. Related to Figure 7

Cell line model	Gene symbol	ENTREZ_ID	Chromosome	Chromosomal start position	Original Base pair	New Base pair	Amino acid change	Mutation type category	Genetic subtype feature
TMD8	BTG2	7832	1	203276274	G	A	G62D	Missense	MCD
TMD8	CD79B	974	17	62006799	A	G	Y196H	Missense	MCD
TMD8	EIF4A2	1974	3	186504979	T	G	F279V	Missense	ST2
TMD8	ETV6	2120	12	11803049	C	G		5UTR	MCD
TMD8	GRHPR	9380	9	37424855	C	T	Q33X	Nonsense	MCD
TMD8	IRF4	3662	6	393318	C	T	H56Y	Missense	MCD
TMD8	MPEG1	219972	11	58979380	G	A	P320L	Missense	MCD
TMD8	MPEG1	219972	11	58978740	G	C	S533R	Missense	MCD
TMD8	MYD88	4615	3	38182641	T	C	L265P	Missense	MCD
TMD8	PIM1	5292	6	37138332	C	T	A85V	Missense	MCD
TMD8	PIM1	5292	6	37138375	G	A	S99S	Synonymous	MCD
TMD8	PIM1	5292	6	37138577	G	C	Q128H	Missense	MCD
TMD8	PIM1	5292	6	37138580	C	A	Y129X	Nonsense	MCD
TMD8	PIM1	5292	6	37138581	C	T	Q130X	Nonsense	MCD
TMD8	PIM1	5292	6	37138804	G	C	E170D	Missense	MCD
TMD8	PIM1	5292	6	37138908	G	A	G174D	Missense	MCD
TMD8	PIM1	5292	6	37139039	C	G	Q218E	Missense	MCD
TMD8	PIM1	5292	6	37139068	G	T	R227S	Missense	MCD
TMD8	PIM1	5292	6	37140860	G	C	L323L	Synonymous	MCD
TMD8	PRDM1	639	6	106536324	G	C	E97D	Missense	MCD
TMD8	PRKCB	5579	16	24183646	T	C	I432T	Missense	BN2
TMD8	TOX	9760	8	59750871	C	G		Splice	MCD
HBL1	ACTB	60	7	5569224	G	A	A22V	Missense	ST2
HBL1	CD79B	974	17	62006798	T	A	Y196F	Missense	MCD
HBL1	DDX3X	1654	X	41206603	G	A	R603Q	Missense	ST2
HBL1	IRF2BP2	359948	1	234744941	G	T	G100G	Synonymous	N1
HBL1	MYD88	4615	3	38182641	T	C	L265P	Missense	MCD
HBL1	NOL9	79707	1	6586772	C	A	C648F	Missense	BN2
HBL1	PIM1	5292	6	37138424	C	T	L116L	Synonymous	MCD
HBL1	TBL1XR1	79718	3	176755958	A	C	N350K	Missense	MCD
HBL1	TP53	7157	17	7578460	A	G	V157A	Missense	A53
OCI-Ly10	ACTG1	71	17	79477761	C	G	E361D	Missense	ST2
OCI-Ly10	BCL2	596	18	60985644	G	A	L86F	Missense	EZB
OCI-Ly10	BCOR	54880	X	39930325	T	A	S1047C	Missense	N1
OCI-Ly10	BTG1	694	12	92539167	C	T	A49T	Missense	MCD
OCI-Ly10	DDX3X	1654	X	41205767	A	G	R503G	Missense	ST2
OCI-Ly10	ETS1	2113	11	128354716	A	G		Splice	BN2
OCI-Ly10	GRHPR	9380	9	37424937	C	T	S60F	Missense	MCD
OCI-Ly10	HLA-B	3106	6	31324049	C	A	E172X	Nonsense	MCD
OCI-Ly10	HLA-C	3107	6	31238865	TCTCCTTCC	D	Y195fs	Frameshift	MCD
OCI-Ly10	IRF2BP2	359948	1	234744673	G	T	P190T	Missense	N1
OCI-Ly10	IRF2BP2	359948	1	234744355	G	A	Q296X	Nonsense	N1
OCI-Ly10	IRF4	3662	6	393243	C	G	Q31E	Missense	MCD
OCI-Ly10	JUNB	3726	19	12902699	C	T	S38S	Synonymous	ST2
OCI-Ly10	MPEG1	219972	11	58980200	G	A	P47S	Missense	MCD
OCI-Ly10	MPEG1	219972	11	58978788	G	A	D517D	Synonymous	MCD
OCI-Ly10	MYD88	4615	3	38182641	T	C	L265P	Missense	MCD
OCI-Ly10	PIM1	5292	6	37138568	G	A	L125L	Synonymous	MCD
OCI-Ly10	PIM1	5292	6	37138649	C	T	N152N	Synonymous	MCD
OCI-Ly10	PIM1	5292	6	37138791	C	T	S166F	Missense	MCD

OCI-Ly10	PIM1	5292	6	37138923	T	G	M179R	Missense	MCD
OCI-Ly10	PIM1	5292	6	37139038	G	A	V217V	Synonymous	MCD
OCI-Ly10	PIM1	5292	6	37139053	C	T	D222D	Synonymous	MCD
OCI-Ly10	PIM2	11040	X	48776062	G	A	T17M	Missense	MCD
OCI-Ly10	PPP1R9B	84687	17	48227539	C	T	K112K	Synonymous	MCD
OCI-Ly10	PRDM1	639	6	106536324	G	C	E97D	Missense	MCD
OCI-Ly10	PRDM1	639	6	106552781	G	D	S249fs	Frameshift	MCD
OCI-Ly10	TOX	9760	8	60031524	GGTG	D	P7fs	Frameshift	MCD
OCI-Ly10	TP53	7157	17	7578406	C	T	R175H	Missense	A53
RIVA	BCL6		3					Translocation	BN2
RIVA	ARID1A	8289	1	27057712	C	T	Q474X	Nonsense	EZB
RIVA	CREBBP	1387	16	3790455	G	A	R1360X	Nonsense	EZB
RIVA	HLA-A	3105	6	29912011	I	T	G245fs	Frameshift	MCD
RIVA	NOTCH2	4853	1	120458147	G	A	R2400X	Nonsense	BN2
RIVA	TP53	7157	17	7578535	T	C	K132R	Missense	A53
RIVA	TP53	7157	17	7577058	C	A	E294X	Nonsense	A53
OCI-Ly1	BCL2							Translocation	BN2
OCI-Ly1	ACTB	60	7	5568262	A	C	I151S	Missense	ST2
OCI-Ly1	B2M	567	15	45003746	T	C	M1?	StartLost	A53
OCI-Ly1	B2M	567	15	45003750	T	G	S2S	Synonymous	A53
OCI-Ly1	B2M	567	15	45007845	T	A	Y98N	Missense	A53
OCI-Ly1	BCL2	596	18	60985508	G	A	A131V	Missense	EZB
OCI-Ly1	BCL2	596	18	60985417	G	A	S161S	Synonymous	EZB
OCI-Ly1	BCL2	596	18	60985340	G	A	T187I	Missense	EZB
OCI-Ly1	EZH2	2146	7	148508728	A	T	Y641N	Missense	EZB
OCI-Ly1	FBXO11	80204	2	48045993	T	C	Y644C	Missense	EZB
OCI-Ly1	FBXO11	80204	2	48035377	I	A	D889fs	Frameshift	EZB
OCI-Ly1	GNA13	10672	17	63052611	A	G	I34T	Missense	EZB
OCI-Ly1	GNA13	10672	17	63010919	A	T	L197Q	Missense	EZB
OCI-Ly1	HIST1H2BK	85236	6	27114367	A	G	F71L	Missense	BN2
OCI-Ly1	HIST1H2BK	85236	6	27114326	G	A	Y84Y	Synonymous	BN2
OCI-Ly1	S1PR2	9294	19	10334849	A	G	C245R	Missense	EZB
OCI-Ly1	S1PR2	9294	19	10334814	G	C	D256E	Missense	EZB
OCI-Ly1	S1PR2	9294	19	10334768	A	T	Y272N	Missense	EZB
OCI-Ly1	TP53	7157	17	7578457	C	T	R158H	Missense	A53
OCI-Ly1	TP53	7157	17	7578404	A	C	C176G	Missense	A53
SUDHL4	BCL2							Translocation	EZB
SUDHL4	ACTB	60	7	5567998	C	T	S239N	Missense	ST2
SUDHL4	BCL2	596	18	60985765	T	C	A45A	Synonymous	EZB
SUDHL4	BCL2	596	18	60985747	G	C	S51S	Synonymous	EZB
SUDHL4	BCL2	596	18	60985725	G	T	P59T	Missense	EZB
SUDHL4	BCL2	596	18	60985549	G	C	S117R	Missense	EZB
SUDHL4	CIITA	4261	16	10997689	C	T	P292S	Missense	EZB
SUDHL4	CIITA	4261	16	11001899	G	D	A851fs	Frameshift	EZB
SUDHL4	DDX3X	1654	X	41205581	A	T	H472L	Missense	ST2
SUDHL4	EZH2	2146	7	148508727	T	G	Y641S	Missense	EZB
SUDHL4	GNAI2	2771	3	50294456	AAG	D	K272del	Indel	EZB
SUDHL4	MEF2B	100271849	19	19260045	T	A	D83V	Missense	EZB
SUDHL4	PIM1	5292	6	37138308	G	A	S77N	Missense	MCD
SUDHL4	PTEN	5728	10	89692918	G	A	M134I	Missense	EZB
SUDHL4	TNFRSF14	8764	1	2489824	T	G	V74G	Missense	EZB
SUDHL4	TP53	7157	17	7577121	G	A	R273C	Missense	A53
WSU-DLCL2	BCL2							Translocation	EZB
WSU-DLCL2	BCL2	596	18	60985960	G	A		5UTR	EZB

WSU-DLCL2	BCL2	596	18	60985931	G	A		5UTR	EZB
WSU-DLCL2	BCL2	596	18	60985869	T	G	N11H	Missense	EZB
WSU-DLCL2	BCL2	596	18	60985861	C	G	E13D	Missense	EZB
WSU-DLCL2	BCL2	596	18	60985725	G	C	P59A	Missense	EZB
WSU-DLCL2	BCL2	596	18	60985705	C	G	P65P	Synonymous	EZB
WSU-DLCL2	BCL2	596	18	60985695	T	G	T69P	Missense	EZB
WSU-DLCL2	BCL2	596	18	60985387	G	A	D171D	Synonymous	EZB
WSU-DLCL2	EZH2	2146	7	148508727	T	A	Y641F	Missense	EZB
WSU-DLCL2	FBXO11	80204	2	48066023	C	A	E188X	Nonsense	EZB
WSU-DLCL2	GNA13	10672	17	63010775	A	G	F245S	Missense	EZB
WSU-DLCL2	ITPKB	3707	1	226924292	C	T	A290T	Missense	ST2
WSU-DLCL2	WEE1	7465	11	9606835	C	T	P440L	Missense	MCD