

Table S1. Mutation Analysis of the S1P2 (EGD5) coding sequences in DLBCL

Tumor Sample	Mutation analysis*	
	Mutations	Polymorphisms [#]
Cell lines		
Ly1	T825C (+/-), C860A (+/-), T906A(+/-)	C1011A (+/+), T1045G (+/+/)
Ly3	-	C1011A (+/-), T1045G (+/+/)
Ly4	-	C1011A (+/+), T1045G (+/+/)
Ly7	-	C1011A (+/-), T1045G (+/+/)
Ly8	-	C1011A (+/-), T1045G (+/+/)
Ly10	-	C1011A (+/-), T1045G (+/+/)
Ly18	-	T1045G (+/+/)
RCK8	-	T1045G (+/+/)
VAL	-	C1011A (+/-), T1045G (+/+/)
SUDHL6	-	C1011A(+/-), T1045G(+/+/)
Primary cases		
424	-	T1045G (+/+/)
1506 [^]	-	T1045G (+/+/)
1508 [^]	-	C122A (+/-),T1045G (+/+/)
1516 [^]	T359A(+/-)	C1011A (+/+), T1045G (+/+/)
1520 [^]	-	C122A (+/-),T1045G (+/+/)
1522 [^]	-	T1045G (+/+/)
1576 [^]	-	C1011A (+/+), T1045G (+/+/)
1580 [^]	-	C1011A, T1045G (+/+/)
1678	-	T1045G (+/+/)
1682	-	T1045G (+/+/)

* Numbering according to Genbank accession No. NM _ 004230

Polymorphisms include changes observed in normal DNA from the same patient, where available (cases indicated by ^)

In bold, missense mutations; +/+, homozygous change; +/-, heterozygous change

Table S2. Mutation analysis of the S1P2 5' sequences in DLBCL

DNA#	Mutations		Polymorphisms**
	N	Position*	
Cell lines			
Ly1	15	T282C, G293A, G349T, G354A, G376A, G430A, C471G, A580G, G582A, T629C, C696T, A922C, C934T, C1039T, C1040A	G1076T
Ly3	0	-	C272T, T1018C, dupTTCA(1271-74)
Ly4	0	-	G1076T(+/+)
Ly7	0	-	C272T
Ly8	3	C772T, G999A, C1034T	C272T(+/+)
Ly10	0	-	G1076T
Ly18	0	-	T1018C, dupTTCA(1271-74)
RCK8	0	-	T1018C, dupTTCA(1271-74)
VAL	0	-	C272T, T1018C, dupTTCA(1271-74)
SUDHL4	1	G183A	G1076T
SUDHL5	2	ΔC(224), C226G	dupTTCA(1271-74)
SUDHL6	0	-	G1076T
SUDHL7	0	-	C272T, dupTTCA(1271-74)
SUDHL8	0	-	G1076T
SUDHL10	1	G138A	dupTTCA(1271-74)
FARAGE	0	-	G1076T
TOLEDO	0	-	G1076T
HT	0	-	G1076T
DB	0	-	dupTTCA(1271-74)
WSU	0	-	G1076T
Primary cases			
1506^	0	-	-
1508^	0	-	-
1516^	5	C349T, G413C, ΔG620, G625A, Δ19bp (626-644)	C272T
1520^	0	-	-
1522^	0	-	-
1576^	0	-	C272T
1580^	0	-	-
1678	0	-	-
424	1	G562A	-

1682	7	C449T, G571C, G582A, C615G, A695G, C871T, C872G	-
2008	0	-	C272T, G1076T
2009	0	-	-
2010	0	-	C272T, G1076T
2011	0	-	C272T
2012	6	C255T, T283A, G638T, A918G, C1033T, C1034T	C272T, C1033G
2013	0	-	-
2014	0	-	C272T, C661A
2015	0	-	G1076T
2016	0	-	C272T, T1018C
2017	0	-	T1018C, dupTTCA(1271-74)
2018	0	-	C272T, G1076T
2019	0	-	C272T(+/+), G1076T
2020	0	-	G1076T
2022	1	C450A	C272T(+/+) G1076T
2024	0	-	-
2025	0	-	G1076T, T1018C, dupTTCA(1271-74)
2026	0	-	T1018C,dupTTCA(1271- 74)(+/+)
2027	1	C727G	C272T(+/+) -
2029	1	G138A	-
2030	0	-	G1076T
2031	0	-	T1018C, dupTTCA(1271- 74)(+/+)
2032	4	G718T, C761G, T736C, G476T	G845A,C1033G
2033	0	-	G1076T
2034	0	-	-
2035	0	-	dupTTCA(1271-74)
2036	1	C846G	C272T,G1076T
2037	2	T923C,T932A	C1033G,C272T,T1076G
2038	1	G1048C	C272T,T1076G,G845A
2039	0	-	-
2040	0	-	C272T G562A, C793T,
2043	0	-	dupTTCA(1271-74)(+/+)
2044	0	-	-
2045	0	-	T1018C,dupTTCA(1271-74)
2046	0	-	T1018C,dupTTCA(1271-74) T1018C(+/+), dupTTCA(1271-74)(+/+)
2047	0	-	

2053	0	-	C272T, G562A,dupTTCA(1271-74)
2054	5	G170A, G302A, G325A,G593C, C1039G	T1018C(+/+),T1076G,dupTT CA(1274-75)
2056	0	-	C272T,T1076G
2057	0	-	T1076G
2058	0	-	T1076G
2059	0	-	-
2060	0	-	C272T
2061	0	-	T1018C, dupTTCA(1271-74)
2062	0	-	C272T(+/)
2063	0	-	G562A, T1018C, dupTTCA(1274-75)
2065	0	-	-
2067	24	C171G, C235A, G254A, C259G, C327G, C419T, C589T, G718A, dup(716-745), T770C, C772A, G790T, C794T, ΔGGTG(839-42), C843A, C846T, A875C, G885C, A893G, G999A, G1015C, C1089T, G1161A, T1240G	C793T, C1034T, G1042C
2069	0	-	T1076G
2073	0	-	C272T, T1018C, dupTTCA(1271-74)
2074	0	-	T1018C, dupTTCA(1271-74)
2075	2	G641A, C666T	C272T(+/)
2076	0	-	T1076G
2078	16	T398G, dup CCAAG(452-456), G467A, C662T, G693A, dup (699-709), T759G, T763G, A818G, C834T, T835A, T841A, G842A, C1024G, C1089G, T1123C	C272T, C1034G, C1039T
2079	4	G582C, A613T, C772T, Δ13bp (1017- 29)	C1034G
2080	0	-	C272T
2082	0	-	T1018C, T1076G, dupTTCA(1271-74)
2085	0	-	C272T
2089	1	C1184T	T1018C, dupTTCA(1271-74)
2091	0	-	C272T
2093	0	-	-
2095	0	-	C272T
2097	0	-	-
2098	1	ΔCA (1257-1258)	-
2099	0	-	T1018C, dupTTCA(1271-74)
2100	4	T347C, G403C, G593T, C859T	C1039T, G1042C
2101	0	-	T1018C(+/),

dupTTCA(1271-74)(++)			
2102	2	G840A, A1038T	T1018C, dupTTCA(1271-74)
2104	0	-	T1018C, dupTTCA(1271-74)
2105	1	C1322T	-
2106	1	G1112T	C272T, G1042C
2107	0	-	C272T, T1076G
2108	0	-	C272T(+/+)
2109	0	-	C272T
2110	1	G639A	C272T
2111	0	-	-
2112	0	-	-

* Position +1 corresponds to the first nucleotide of GenBank accession No. NM_004230

** Polymorphisms included known germline variants, changes present in matched normal DNA from the same patient (samples marked by ^), and changes recurring in more than 2 patients.

Δ, deletion; dup, duplication; +/-, homozygous change

Table S3. Mutation analysis of IgV, BCL6 and S1P2 sequences in normal B subpopulations

Gene (fragment size)	Cell Population	N of clones analyzed	N of clones mutated	N of bp analyzed	N of mutated bases	Frequency of mutations	p value* (X2 test)
IgVH (~250bp)	naïve	20	1	5000	4	0.0800	
	GC	19	18	4750	242	5.0947	0.0000
	Fibroblasts	na	na	na	na	na	
BCL-6^ (750bp)	naïve	13	2	9750	2	0.021	
	GC	12	5	9000	17	0.189	0.0003; 0.0000
	Fibroblasts	20	0	15000	0	0.000	
S1P2 (1252 bp)	naïve	75	6	93900	6	0.0064	
	GC	122	14	152744	20	0.0131	0.115; 0.445 (ns)
	Fibroblasts	82	10	102500	11	0.0107	

*Calculated by comparing GC cells to naïve B cells and fibroblasts, respectively, which were both used as controls for the Pfu polymerase rate, since the SHM mechanism is not active in these populations. ns, not significant

^ data from Pasqualucci et al., Nature 2001