Supplementary Materials

Supplementary Table 1. eQTL analysis for rs9831894 and rs6773363

GWAS SNP	SNPs in LD	#Chr	Position	eQTL Cluster (GTEx)
rs9831894	rs9831894	3	121800487	CD86;testis,
rs9831894	rs9828869	3	121748360	CD86;testis,
rs9831894	rs1962532	3	121751531	0
rs9831894	rs527769191	3	121759584	0
rs9831894	rs62269210	3	121760603	0
rs9831894	rs62269211	3	121760639	0
rs9831894	rs28876421	3	121764257	0
rs9831894	rs2255214	3	121770539	0
rs9831894	rs17203397	3	121781426	0
rs9831894	rs6786977	3	121782806	0
rs9831894	rs2681411	3	121785631	0
rs9831894	rs4308217	3	121793187	0
rs9831894	rs35187120	3	121796389	0
rs9831894	rs11717893	3	121806825	0
rs9831894	rs9836399	3	121811811	CD86;testis,
rs9831894	rs9289187	3	121815461	CD86;testis,
rs9831894	rs2681416	3	121817613	CD86;testis,
rs9831894	rs2332096	3	121821143	CD86;testis,
rs9831894	rs1915092	3	121823375	0
rs9831894	rs2262495	3	121841801	0
rs9831894	rs2681401	3	121843127	0
rs9831894	rs2715272	3	121843241	0
rs6773363	rs6773363	3	27793632	0
rs6773363	rs2371108	3	27757018	RP11-222K16.2;Esophagus - Muscularis,
rs6773363	rs530939637	3	27757867	0
rs6773363	rs2371109	3	27758274	0
rs6773363	rs2887944	3	27758275	0
rs6773363	rs9866625	3	27759338	0
rs6773363	rs56306375	3	27760524	0
rs6773363	rs368178421	3	27763427	0
rs6773363	rs3806624	3	27764623	0
rs6773363	rs3840481	3	27765595	0
rs6773363	rs34269949	3	27769551	0
rs6773363	rs1353286	3	27772014	0
rs6773363	rs5847475	3	27772346	0
rs6773363	rs549698945	3	27774208	0
rs6773363	rs6801231	3	27774245	0
rs6773363	rs34860778	3	27776242	0
rs6773363	rs9880772	3	27777779	0
rs6773363	rs12635205	3	27779362	0
rs6773363	rs34024889	3	27780638	0
rs6773363	rs13327021	3	27783015	0
rs6773363	rs4680838	3	27783476	0
rs6773363	rs9310852	3	27784997	0
rs6773363	rs145548114	3	27786022	0
rs6773363	rs11129295	3	27788780	0
rs6773363	rs12330493	3	27791614	0

rs6773363	rs6806966	3	27793862	0
rs6773363	rs13093489	3	27794907	0
rs6773363	rs7371953	3	27795157	0
rs6773363	rs12497690	3	27795397	0
rs6773363	rs113106250	3	27797023	0
rs6773363	rs199900263	3	27797030	0
rs6773363	rs12493245	3	27797810	0
rs6773363	rs4680843	3	27798071	0
rs6773363	rs4680844	3	27798172	0
rs6773363	rs9860067	3	27798275	0
rs6773363	rs4680845	3	27799034	0
rs6773363	rs17020283	3	27799105	0
rs6773363	rs572217447	3	27799483	0
rs6773363	rs572217447	3	27799483	0
rs6773363	rs9862284	3	27800325	0
rs6773363	rs6795177	3	27800734	0
rs6773363	rs17020287	3	27801879	0
rs6773363	rs60047326	3	27801932	0
rs6773363	rs17020288	3	27801959	0
rs6773363	rs9814971	3	27802163	0
rs6773363	rs9872851	3	27802302	0
rs6773363	rs9815263	3	27802364	0
rs6773363	rs9815419	3	27802488	0
rs6773363	rs56143766	3	27802504	0
rs6773363	rs9816235	3	27803017	0
rs6773363	rs4680730	3	27804731	0
rs6773363	rs4680851	3	27805177	0
rs6773363	rs4680852	3	27805204	0
rs6773363	rs5847478	3	27805726	0
rs6773363	rs2371150	3	27806147	0
rs6773363	rs4450763	3	27806852	0
rs6773363	rs4522733	3	27807006	0
rs6773363	rs4371473	3	27807063	0
rs6773363	rs12497514	3	27807665	0
rs6773363	rs55998928	3	27807730	0
rs6773363	rs11711117	3	27809342	0
rs6773363	rs200433332	3	27809401	0
rs6773363	rs1000361	3	27810319	0
rs6773363	rs9868208	3	27810888	0
rs6773363	rs5012869	3	27811220	0
rs6773363	rs62253101	3	27811637	0
rs6773363	rs9878524	3	27812457	0
rs6773363	rs9878720	3	27812594	0
rs6773363	rs9823274	3	27812666	0
rs6773363	rs4680861	3	27812802	0
rs6773363	rs4279053	3	27813319	0
rs6773363	rs11915329	3	27814628	0
rs6773363	rs62253104	3	27815285	0
rs6773363	rs201123277	3	27815370	0
rs6773363	rs6551228	3	27815535	0
rs6773363	rs10049011	3	27815942	0
rs6773363	rs13083047	3	27817959	0
rs6773363	rs9872426	3	27818453	0
rs6773363	rs7628641	3	27818887	0

rs6773363	rs7628648	3	27818893	0
rs6773363	rs7621606	3	27819783	0
rs6773363	rs12633504	3	27819955	0
rs6773363	rs9838021	3	27820995	0
rs6773363	rs60294567	3	27821359	0
rs6773363	rs7646728	3	27822171	0
rs6773363	rs7647094	3	27822530	0
rs6773363	rs545045422	3	27823712	0
rs6773363	rs545045422	3	27823712	0
rs6773363	rs7627710	3	27823887	0
rs6773363	rs11710954	3	27824022	0
rs6773363	rs7653004	3	27824089	0
rs6773363	rs11711718	3	27824134	0
rs6773363	rs57583032	3	27824709	0
rs6773363	rs11129297	3	27825386	0
rs6773363	rs7633952	3	27825912	0
rs6773363	rs7626965	3	27826752	0
rs6773363	rs4491854	3	27826932	0
rs6773363	rs34089683	3	27827398	0
rs6773363	rs9817112	3	27827733	0
rs6773363	rs4680868	3	27829161	RP11-222K16.1;nerve - Tibial,
rs6773363	rs11925920	3	27829393	0
rs6773363	rs561194513	3	27829950	0
rs6773363	rs11712606	3	27830379	0
rs6773363	rs543125886	3	27830535	0
rs6773363	rs556526857	3	27830865	0
rs6773363	rs60579266	3	27831914	0
rs6773363	rs6551229	3	27831978	0

eQTL, expression quantitative trait loci; GTEx, Genotype-Tissue Expression

Supplementary Table 2. Summary of the participating studies and genotyping methods in discovery.

Study	Study name	No. cases	No. controls	Design	Description of the study	Genotyping platform	Genotype calling algorithm	Imputation Software
NCI-SEER	National Cancer Institute-Surveillance, Epidemiology, and End Results Interdisciplinary Case-Control Study of Non-Hodgkin Lymphoma	2661	6221	Pooling of case-control studies, nested case-control studies from cohorts	8 Population-based case-control studies. 5 clinic or hospital-based or mixed case-control studies. 9 cohort studies.	Illumina OmniExpress	BeadStudio (GenCall)	IMPUTE2
GELA	Treatment program of DLBCL patients from the Groupe d'Etude des Lymphomes de l'Adulte (GELA) consisting in LNH03-1B, 2B, 3B, 39B, 6B and 7B.	549	525	Cases from randomized clinical trial with population-based controls from EPIC	Patients recruited in the study that consented to DNA analysis. Controls were from the EPIC cohort	Illumina HumanHap 610K (cases) Illumina HumanHap 610K and 660W BeadStudio (controls)	BeadStudio (GenCall)	IMPUTE2
MAYO	Mayo Clinic Case-Control Study of NHL and CLL and Iowa-Mayo SPORE Molecular Epidemiology Resource	393	172	Clinic-based case-control study and clinic-based registry	Consecutive patients with newly diagnosed, histologically-confirmed non-Hodgkin lymphoma (excluding HIV-infected cases) who were residents of Minnesota, Iowa or Wisconsin. Controls were selected from patients seen in the general medicine clinics at Mayo with a pre-scheduled general medical examination, frequency on age, sex, and geographic region	Illumina HumanHap 660W	BeadStudio (GenCall)	IMPUTE2

UCSF 2	Molecular Epidemiology of non-Hodgkin lymphoma	254	748	Population-based case-control study	RCA/SEER Incident NHL diagnosis for patients diagnosed in hospitals in 6 San Francisco Bay Area Counties and who were residents of the Bay Area at the time of diagnosis. Controls ascertained through RDD were frequency matched to cases on age in 5-year groups, sex and county of residence; Random sampling of CMS lists for person residing in the same 6 Bay Area counties were used to supplement recruitment of controls aged 65+.	Illumina HumanCNV370-Duo	BeadStudio (GenCall)	IMPUTE2

Supplementary Table 3. Summary of the replication studies conducting de novo genotyping for this study.

Study	**Study name**	**Design**	**Description of the study**	**SNP**	**Genotyping platform**	**Quality Control**	**Final No. Cases**	**Final No. Controls**
MAYO	Iowa-Mayo SPORE Molecular Epidemiology Resource Mayo Clinic Case-Control Study of NHL	Clinic-based case registry and Clinic-based case-control study	See supplementary Table 1.	rs9831894	Sequenom	99% concordance on 48 duplicates	782	1053
				rs6773363	Taqman	100% concordance on 11 duplicates	659	1059
MDA	MD Anderson lymphoma case-control study	Case-control	Cases from MD Anderson Cancer Center. Controls from Kelsey Seybold Clinics.	rs9831894	Taqman	100% concordance on 6% replicates + 2 CEPH samples	371	370
MSKCC	Memorial-Sloan Kettering Lymphoproliferative disorders Study	Hospital-based case-study and NYCP controls	Cases were hospital clinic based ascertainment in a tertiary referral center. NYCP controls from same geographic area.	rs9831894	Sequenom	100% concordance on 24 duplicates	653	148

Supplementary Table 4. Public epigenetic data from B cells, lymphoma cell lines and patients

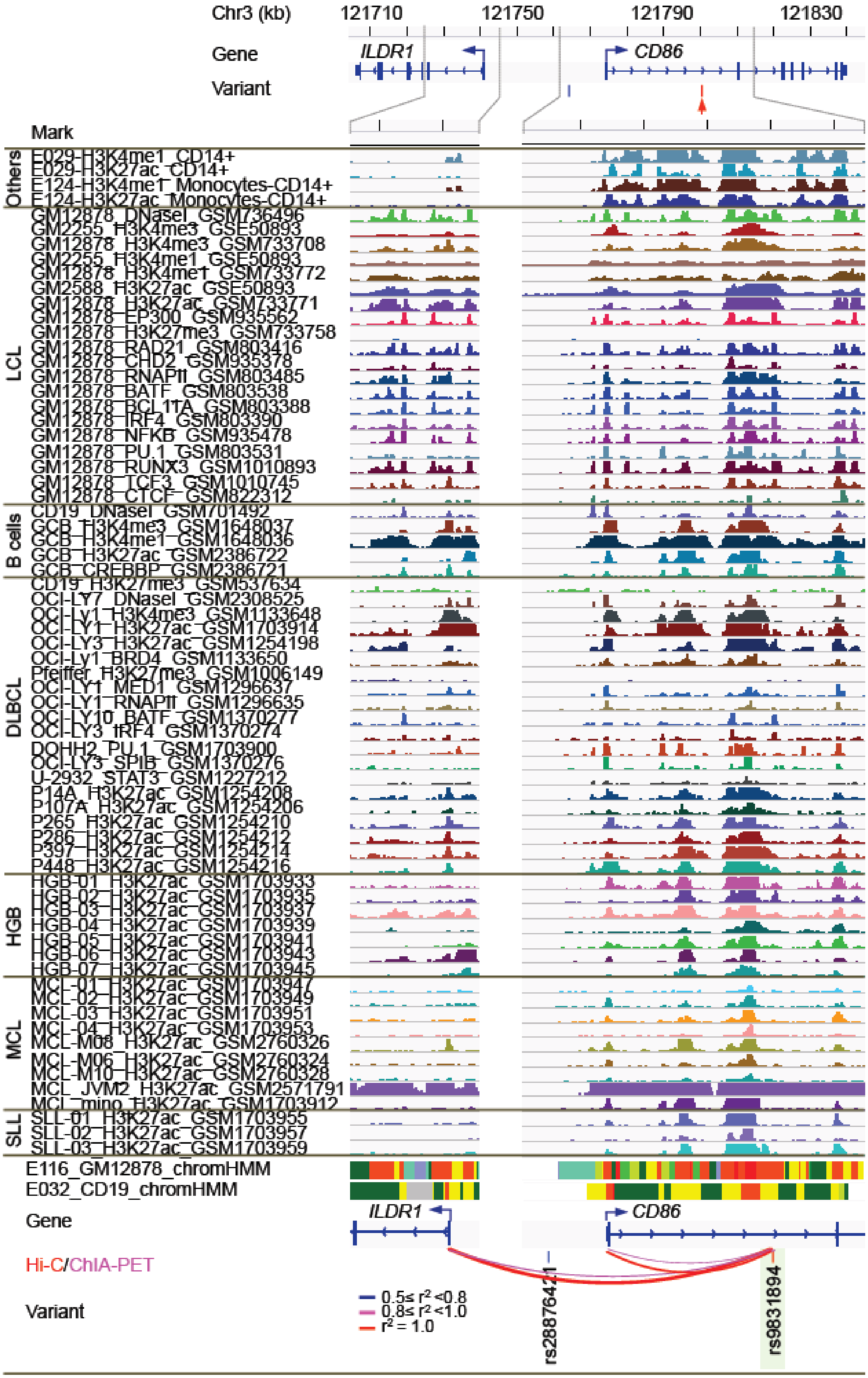
Sample	Mark	No. sample	ID	Format	Ref.
B cells	CREBBP	1	GSM2386721	sra	(1)
B cells	DNaseI	1	GSM701492	sra	(2)
B cells	H3K27ac	1	GSM1027287	sra	(2)
B cells	H3K27ac	1	GSM1519135	sra	(3)
B cells	H3K27ac	1	GSM1519136	sra	(3)
B cells	H3K27ac	1	GSM1519134	sra	(3)
B cells	H3K27ac	1	GSM1519137	sra	(3)
B cells	H3K27ac	1	GSM2386722	sra	(1)
B cells	H3K27ac	1	GSM2386720	sra	(1)
B cells	H3K27me3	1	GSM971339	sra	(4)
B cells	H3K27me3	1	GSM537634	tagalign	(2)
B cells	H3K27me3	1	GSM537638	tagalign	(2)
B cells	H3K4me1	1	GSM916053	sra	(2)
B cells	H3K4me1	1	GSM1648036	sra	(5)
B cells	H3K4me3	1	GSM537622	tagalign	(2)
B cells	H3K4me3	1	GSM1648037	sra	(5)
DLBCL cell line OCI-LY1	BRD4	1	GSM1296634	sra	
DLBCL cell line OCI-Ly1	BRD4	1	GSM1133650	sra	(6)
DLBCL cell line OCI-LY7	DHS	1	GSM2308524	sra	(2)
DLBCL cell line OCI-LY7	DHS	1	GSM2308525	sra	(2)
DLBCL cell line OCI-Ly1	H3K27ac	1	GSM1133646	sra	(6)
DLBCL cell line DHL6	H3K27ac	1	GSM1254194	sra	(6)
DLBCL cell line HBL1	H3K27ac	1	GSM1254196	sra	(6)
DLBCL cell line OCI-LY3	H3K27ac	1	GSM1254198	sra	(6)
DLBCL cell line OCI-LY4	H3K27ac	1	GSM1254200	sra	(6)
DLBCL cell line TOLEDO	H3K27ac	1	GSM1254202	sra	(6)
DLBCL cell line DOHH2	H3K27ac	1	GSM1703895	sra	(7)
DLBCL cell line KARPAS-422	H3K27ac	1	GSM1703909	sra	(7)
DLBCL cell line OCI-LY1	H3K27ac	1	GSM1703914	sra	(7)
DLBCL cell line OCI-LY7	H3K27ac	1	GSM1703918	sra	(7)
DLBCL cell line PFEIFFER	H3K27ac	1	GSM1703923	sra	(7)
DLBCL cell line SU-DHL4	H3K27ac	1	GSM1703927	sra	(7)
DLBCL cell line TOLEDO	H3K27ac	1	GSM1703931	sra	(7)
DLBCL cell line OCI-LY1	H3K27ac	1	GSM2037789	sra	(8)
DLBCL cell line Pfeiffer	H3K27me3	1	GSM1006149	sra	(9)
DLBCL cell line KARPAS-422	H3K27me3	1	GSM1006151	sra	(9)
DLBCL cell line WSU-DLCL2	H3K27me3	1	GSM1006153	sra	(9)
DLBCL cell line OCI-Ly1	H3K4me3	1	GSM1572767	sra	(10)
DLBCL cell line OCI-Ly1	H3K4me3	1	GSM1133648	sra	(6)
DLBCL cell line OCI-LY1	MED1	1	GSM1296637	sra	
DLBCL cell line DB	MEF2B	1	GSM1703891	sra	(7)
DLBCL cell line DB	MEF2B	1	GSM1703892	sra	(7)
DLBCL cell line DOHH2	MEF2B	1	GSM1703898	sra	(7)
DLBCL cell line KARPAS-422	MEF2B	1	GSM1703911	sra	(7)
DLBCL cell line SU-DHL4	MEF2B	1	GSM1703929	sra	(7)
DLBCL cell line SU-DHL4	MEF2B	1	GSM1703930	sra	(7)
DLBCL cell line OCI-LY1	RNAPII	1	GSM1296635	sra	
DLBCL cell line OCI-Ly1	RNAPII	1	GSM1133652	sra	(6)

DLBCL cell line DOHH2	BACH2	1	GSM1703893	sra	(7)
DLBCL cell line OCI-LY7	BACH2	1	GSM1703916	sra	(7)
DLBCL cell line OCI-LY3	BATF	1	GSM1370272	sra	(11)
DLBCL cell line OCI-LY10	BATF	1	GSM1370277	sra	(11)
DLBCL cell line OCI-LY1	BCL6	1	GSM2754202	sra	(12)
DLBCL cell line OCI-LY1	BCL6	1	GSM2754203	sra	(12)
DLBCL cell line OCI-Ly1	BCL6	1	GSM1133657	sra	(6)
DLBCL cell line OCI-Ly1	E2F1	1	GSM1133656	sra	(6)
DLBCL cell line OCI-LY3	IRF4	1	GSM1370274	sra	(11)
DLBCL cell line OCI-LY10	IRF4	1	GSM1370279	sra	(11)
DLBCL cell line DOHH2	PAX5	1	GSM1703899	sra	(7)
DLBCL cell line OCI-LY7	PAX5	1	GSM1703921	sra	(7)
DLBCL cell line DOHH2	PU.1	1	GSM1703900	sra	(7)
DLBCL cell line OCI-LY7	PU.1	1	GSM1703922	sra	(7)
DLBCL cell line OCI-LY3	SPIB	1	GSM1370276	sra	(11)
DLBCL cell line OCI-LY10	SPIB	1	GSM1370281	sra	(11)
DLBCL cell line OCI-Ly19	STAT3	1	GSM1227205	sra	(13)
DLBCL cell line OCI-Ly3	STAT3	1	GSM1227206	sra	(13)
DLBCL cell line OCI-Ly7	STAT3	1	GSM1227207	sra	(13)
DLBCL cell line SU-DHL10	STAT3	1	GSM1227208	sra	(13)
DLBCL cell line SU-DHL2	STAT3	1	GSM1227209	sra	(13)
DLBCL cell line SU-DHL4	STAT3	1	GSM1227210	sra	(13)
DLBCL patients	H3K27ac	6	GSM1254206,GSM1254208,GSM1254210,GSM1254212,GSM1254214,GSM1254216	sra	(6)
HGB patients	H3K27ac	7	GSM1703933,GSM1703935,GSM1703937,GSM1703939,GSM1703941,GSM1703943,GSM1703945	tdf	(7)
MCL cell lines GRANTA-519	H3K27ac	1	GSM1703901	tdf	(7)
MCL cell lines JEKO-1	H3K27ac	1	GSM1703903	tdf	(7)
MCL cell lines MINO	H3K27ac	1	GSM1703912	tdf	(7)
MCL cell lines REC-1	H3K27ac	1	GSM1703925	tdf	(7)
MCL cell lines HBL2	H3K27ac	1	GSM2571789	tdf	(14)
MCL cell lines JVM2	H3K27ac	1	GSM2571791	tdf	(14)
MCL cell lines MAVER-1	H3K27ac	1	GSM2571793	tdf	(14)
MCL cell lines SP-49	H3K27ac	1	GSM2571828	tdf	(14)
MCL cell lines Z138	H3K27ac	1	GSM2571832	tdf	(14)
MCL patients	H3K27ac	4	GSM1703947,GSM1703949,GSM1703951,GSM1703953	tdf	(7)
MCL patients	H3K27ac	3	GSM2760324,GSM2760326,GSM2760328	tdf	(14)
SLL patients	H3K27ac	3	GSM1703955,GSM1703957,GSM1703959	tdf	(7)
FL patients	H3K27ac	9	GSM1523532,GSM1523535,GSM1523544,GSM1523550,GSM1523556,GSM1523559,GSM1523565, GSM1523571,GSM1523573	bigwig	(15)

Except DNase I hypersensitive site (DHS), other marks were assayed by ChIP-seq. FL, follicular lymphoma; HGB, high-grade B-cell lymphoma; MCL, mantle cell lymphoma; SLL, small lymphocytic lymphoma.

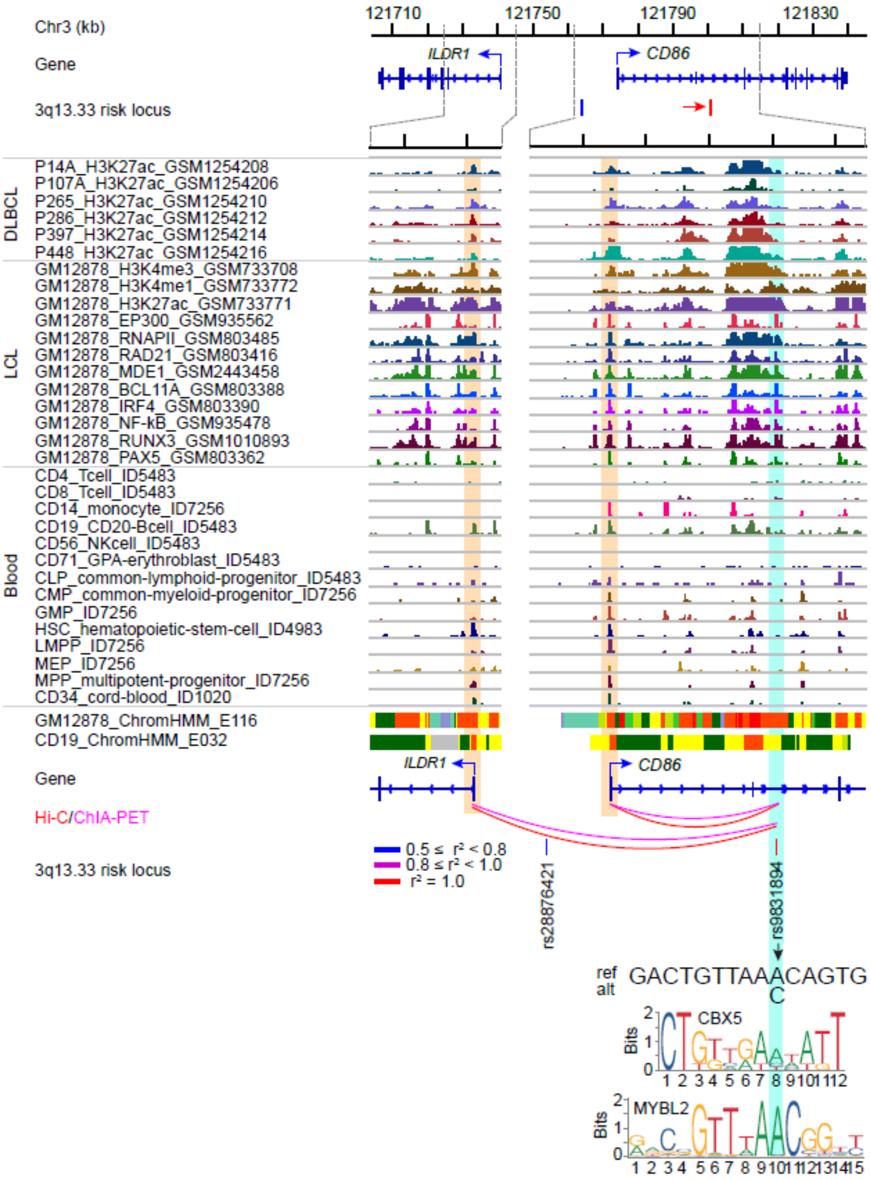
Supplementary Figure 1. Epigenetic features of rs9831894 at independent loci associated with the risk of DLBCL. Epigenetic features at the 3q13.33 risk locus. Tag SNP rs9831894 has a single linked variant ($r^2 \geq 0.5$). Only rs9831894 was shown for its overlap with epigenetic marks, while the other one was in a region without the enrichment of the assayed epigenetic marks. In particular, The GWAS SNP rs9831894 is in a super-enhancer. The region spanning rs9831894 showed looping interactions with the promoters of *CD86* and *ILDR1*. E029, primary monocytes from peripheral blood (CD14+); E124, monocytes CD14+ RO01746 primary cells.

Chromatin state
🟥 1 Active TSS
🟧 2 Flanking active TSS
🟩 3 Transcription at gene 5′ and 3′
🟩 4 Strong transcription
🟩 5 Weak transcription
🟨 6 Genic enhancers
🟨 7 Enhancers
🟩 8 ZNF genes + repeats
🟦 9 Heterochromatin
🟥 10 Bivalent/poised TSS
🟧 11 Flanking bivalent TSS/Enh
🟫 12 Bivalent enhancer
⬛ 13 Repressed Polycomb
⬜ 14 Weak repressed Polycomb
15 Quiescent/low



Chr3 (kb)
121710
121750
121790
121830
Gene
ILDR1
CD86
Variant
Mark
Others
E029-H3K4me1_CD14+
E029-H3K27ac_CD14+
E124-H3K4me1_Monocytes-CD14+
E124-H3K27ac_Monocytes-CD14+
LCL
GM12878_DNaseI_GSM736496
GM2255_H3K4me3_GSE50893
GM12878_H3K4me3_GSM733708
GM2255_H3K4me1_GSE50893
GM12878_H3K4me1_GSM733772
GM2588_H3K27ac_GSE50893
GM12878_H3K27ac_GSM733771
GM12878_EP300_GSM935562
GM12878_H3K27me3_GSM733758
GM12878_RAD21_GSM803416
GM12878_CHD2_GSM935378
GM12878_RNAPII_GSM803485
GM12878_BATF_GSM803538
GM12878_BCL11A_GSM803388
GM12878_IRF4_GSM803390
GM12878_NFKB_GSM935478
GM12878_PU.1_GSM803531
GM12878_RUNX3_GSM1010893
GM12878_TCF3_GSM1010745
GM12878_CTCF_GSM822312
B cells
CD19_DNaseI_GSM701492
GCB_H3K4me3_GSM1648037
GCB_H3K4me1_GSM1648036
GCB_H3K27ac_GSM2386722
GCB_CREBBP_GSM2386721
DLBCL
CD19_H3K27me3_GSM537634
OCI-LY7_DNaseI_GSM2308525
OCI-Ly1_H3K4me3_GSM1133648
OCI-LY1_H3K27ac_GSM1703914
OCI-LY3_H3K27ac_GSM1254198
OCI-Ly1_BRD4_GSM1133650
Pfeiffer_H3K27me3_GSM1006149
OCI-LY1_MED1_GSM1296637
OCI-LY1_RNAPII_GSM1296635
OCI-LY10_BATF_GSM1370277
OCI-LY3_IRF4_GSM1370274
DOHH2_PU.1_GSM1703900
OCI-LY3_SPIB_GSM1370276
U-2932_STAT3_GSM1227212
P14A_H3K27ac_GSM1254208
P107A_H3K27ac_GSM1254206
P265_H3K27ac_GSM1254210
P286_H3K27ac_GSM1254212
P397_H3K27ac_GSM1254214
P448_H3K27ac_GSM1254216
HGB
HGB-01_H3K27ac_GSM1703933
HGB-02_H3K27ac_GSM1703935
HGB-03_H3K27ac_GSM1703937
HGB-04_H3K27ac_GSM1703939
HGB-05_H3K27ac_GSM1703941
HGB-06_H3K27ac_GSM1703943
HGB-07_H3K27ac_GSM1703945
MCL
MCL-01_H3K27ac_GSM1703947
MCL-02_H3K27ac_GSM1703949
MCL-03_H3K27ac_GSM1703951
MCL-04_H3K27ac_GSM1703953
MCL-M08_H3K27ac_GSM2760326
MCL-M06_H3K27ac_GSM2760324
MCL-M10_H3K27ac_GSM2760328
MCL_JVM2_H3K27ac_GSM2571791
MCL_mino_H3K27ac_GSM1703912
SLL
SLL-01_H3K27ac_GSM1703955
SLL-02_H3K27ac_GSM1703957
SLL-03_H3K27ac_GSM1703959
E116_GM12878_chromHMM
E032_CD19_chromHMM
Gene
ILDR1
CD86
Hi-C/ChIA-PET
Variant
0.5≤ r² <0.8
0.8≤ r² <1.0
r² = 1.0
rs28876421
rs9831894

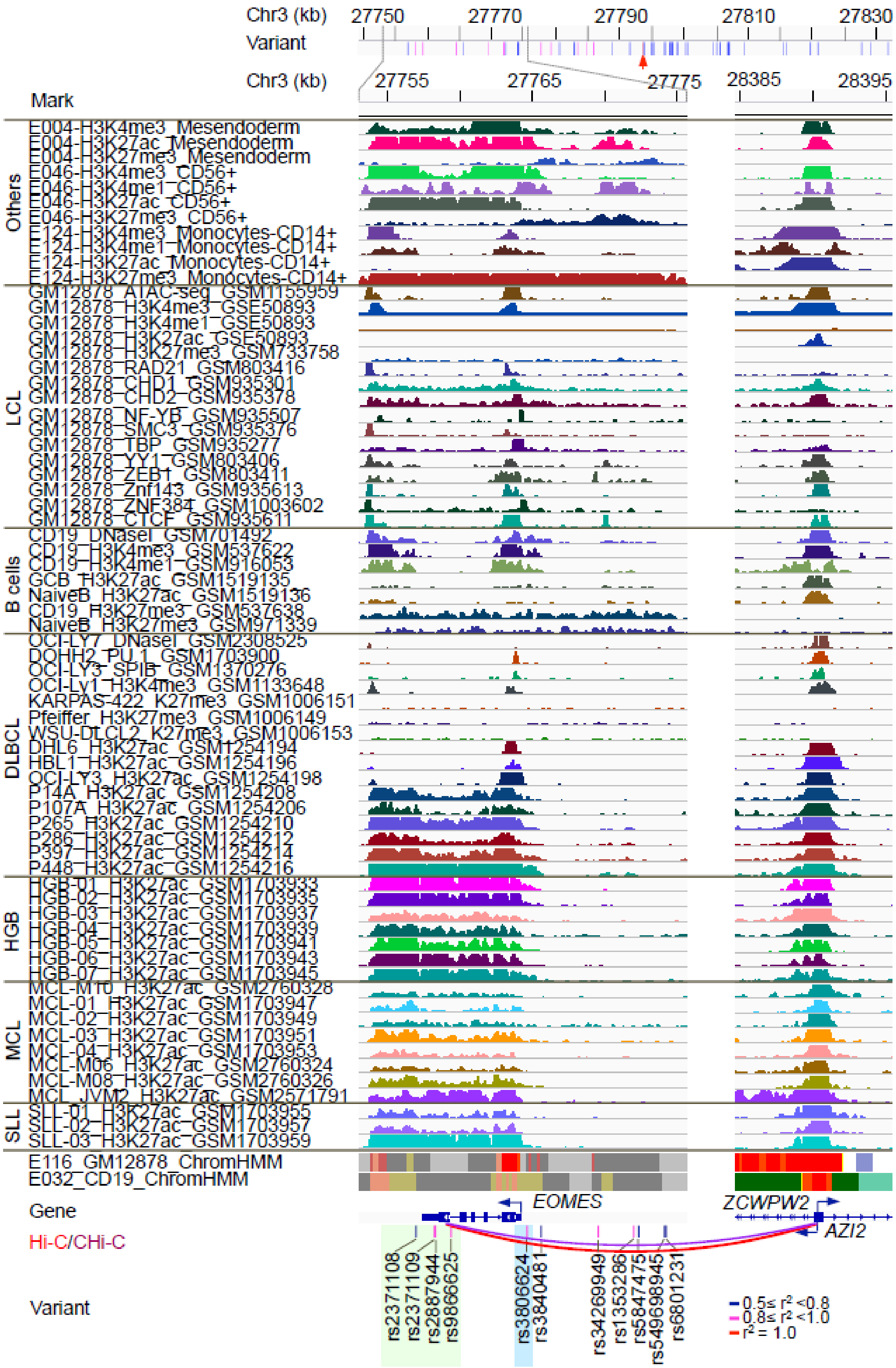
Supplementary Figure 2. Regulatory variant at rs9831894-tagged DLBCL risk locus on 3q13.33. The ATAC-seq data for the 14 blood cell types was downloaded from NCBI GEO under the accession GSE74912. The enhancer carrying tag SNP rs9831894 was found to be transcribed into eRNA in GM12878 (highlighted). In GM12878, this enhancer, like the *ILDR1* and *CD86* promoter it interacted with, showed the binding of RNAPII, P300, RAD21, MED1, and multiple TFs, as typically seen for transcribed enhancer. GMP, granulocyte macrophage progenitor; LMPP, lymphoid primed multipotent progenitor; MEP, megakaryocyte erythroid progenitor. See Supplementary Figure 1 legend for additional information.



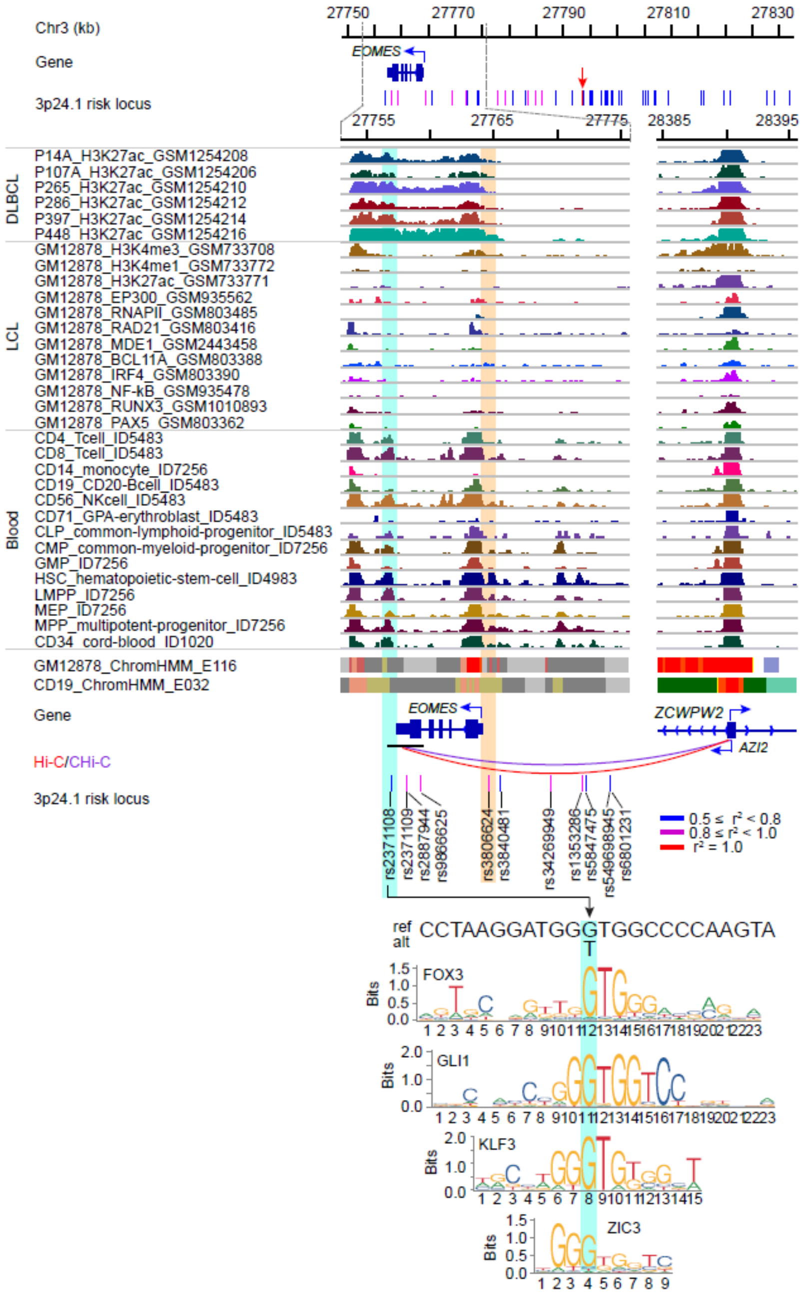
Supplementary Figure 3. Epigenetic features of rs6773363 at independent loci associated with the risk of DLBCL. Epigenetic features at the 3p24.1 risk locus. This 74,961-bp risk locus contains 51 variants ($r^2 \geq 0.5$), with red arrow indicating the tag SNP rs6773363. For simplicity, only 11 of the variants were displayed (bottom panel) for the association with epigenetic profiles. Forty-three of the variants starting from rs34269949, including the tag SNP, did not overlap any of the active marks assayed in B-cells and DLBCL. Four variants between rs2371108 to rs9866625 ($r^2 > 0.7$) overlapped an 11.6-kb super-enhancer. Based on the Hi-C data in GM12878 and capture Hi-C data from 17 blood cell types, this region showed strong long-range interactions with the promoters of *AZI2* and *ZCWPW2*. Rs3806624 (highlighted in light blue), located 417 bp upstream of the transcription start site of *EOMES*, is in the edge of the super-enhancer. E004, H1 BMP4 derived mesendoderm cultured cells; E046, primary natural killer cells from peripheral blood (CD56+); E124, monocytes-CD14+ RO01746 primary cells.

Chromatin state

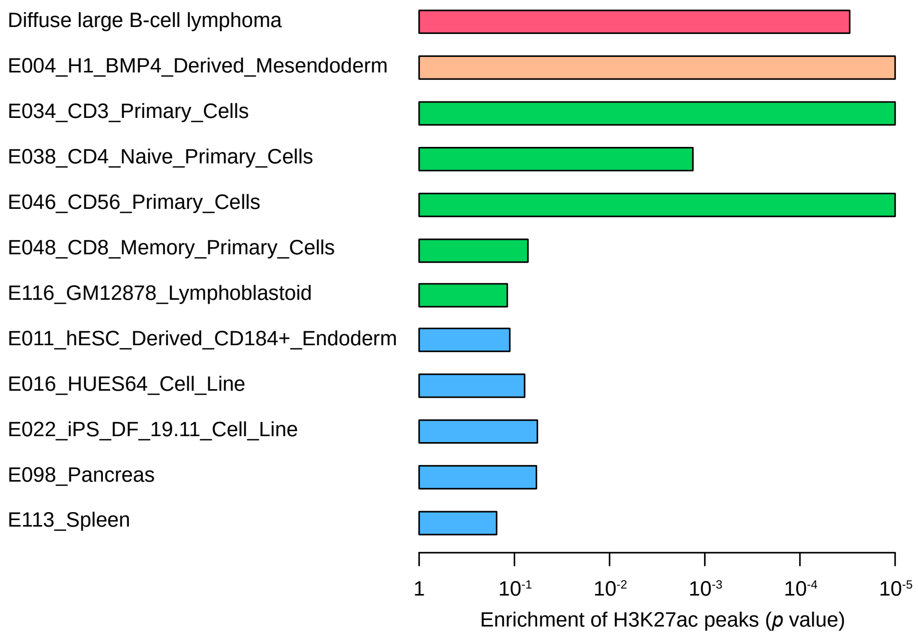
🟥 1 Active TSS
🟧 2 Flanking active TSS
🟩 3 Transcription at gene 5′ and 3′
🟩 4 Strong transcription
🟩 5 Weak transcription
🟩 6 Genic enhancers
🟨 7 Enhancers
🟩 8 ZNF genes + repeats
🟪 9 Heterochromatin
🟥 10 Bivalent/poised TSS
🟧 11 Flanking bivalent TSS/Enh
🟨 12 Bivalent enhancer
⬛ 13 Repressed Polycomb
⬜ 14 Weak repressed Polycomb
15 Quiescent/low


Chr3 (kb)
27750
27770
27790
27810
27830
Variant
Chr3 (kb)
27755
27765
27775
28385
28395
Mark
Others
E004-H3K4me3_Mesendoderm
E004-H3K27ac_Mesendoderm
E004-H3K27me3_Mesendoderm
E046-H3K4me3_CD56+
E046-H3K4me1_CD56+
E046-H3K27ac_CD56+
E046-H3K27me3_CD56+
E124-H3K4me3_Monocytes-CD14+
E124-H3K4me1_Monocytes-CD14+
E124-H3K27ac_Monocytes-CD14+
E124-H3K27me3_Monocytes-CD14+
LCL
GM12878_ATAC-seq_GSM1155959
GM12878_H3K4me3_GSE50893
GM12878_H3K4me1_GSE50893
GM12878_H3K27ac_GSE50893
GM12878_H3K27me3_GSM733758
GM12878_RAD21_GSM803416
GM12878_CHD1_GSM935301
GM12878_CHD2_GSM935378
GM12878_NF-YB_GSM935507
GM12878_SMC3_GSM935376
GM12878_TBP_GSM935277
GM12878_YY1_GSM803406
GM12878_ZEB1_GSM803411
GM12878_Znf143_GSM935613
GM12878_ZNF384_GSM1003602
GM12878_CTCF_GSM935611
B cells
CD19_DNaseI_GSM701492
CD19_H3K4me3_GSM537622
CD19_H3K4me1_GSM916053
GCB_H3K27ac_GSM1519135
NaiveB_H3K27ac_GSM1519136
CD19_H3K27me3_GSM537638
NaiveB_H3K27me3_GSM971339
DLBCL
OCI-LY7_DNaseI_GSM2308525
DOHH2_PU.1_GSM1703900
OCI-LY3_SPIB_GSM1370276
OCI-Ly1_H3K4me3_GSM1133648
KARPAS-422_K27me3_GSM1006151
Pfeiffer_H3K27me3_GSM1006149
WSU-DLCL2_K27me3_GSM1006153
DHL6_H3K27ac_GSM1254194
HBL1_H3K27ac_GSM1254196
OCI-LY3_H3K27ac_GSM1254198
P14A_H3K27ac_GSM1254208
P107A_H3K27ac_GSM1254206
P265_H3K27ac_GSM1254210
P286_H3K27ac_GSM1254212
P397_H3K27ac_GSM1254214
P448_H3K27ac_GSM1254216
HGB
HGB-01_H3K27ac_GSM1703933
HGB-02_H3K27ac_GSM1703935
HGB-03_H3K27ac_GSM1703937
HGB-04_H3K27ac_GSM1703939
HGB-05_H3K27ac_GSM1703941
HGB-06_H3K27ac_GSM1703943
HGB-07_H3K27ac_GSM1703945
MCL
MCL-M10_H3K27ac_GSM2760328
MCL-01_H3K27ac_GSM1703947
MCL-02_H3K27ac_GSM1703949
MCL-03_H3K27ac_GSM1703951
MCL-04_H3K27ac_GSM1703953
MCL-M06_H3K27ac_GSM2760324
MCL-M08_H3K27ac_GSM2760326
MCL_JVM2_H3K27ac_GSM2571791
SLL
SLL-01_H3K27ac_GSM1703955
SLL-02_H3K27ac_GSM1703957
SLL-03_H3K27ac_GSM1703959
E116_GM12878_ChromHMM
E032_CD19_ChromHMM
Gene
EOMES
ZCWPW2
AZI2
Hi-C/CHi-C
Variant
rs2371108
rs2371109
rs2887944
rs9866625
rs3806624
rs3840481
rs34269949
rs1353286
rs5847475
rs549698945
rs6801231
0.5≤ r² <0.8
0.8≤ r² <1.0
r² = 1.0

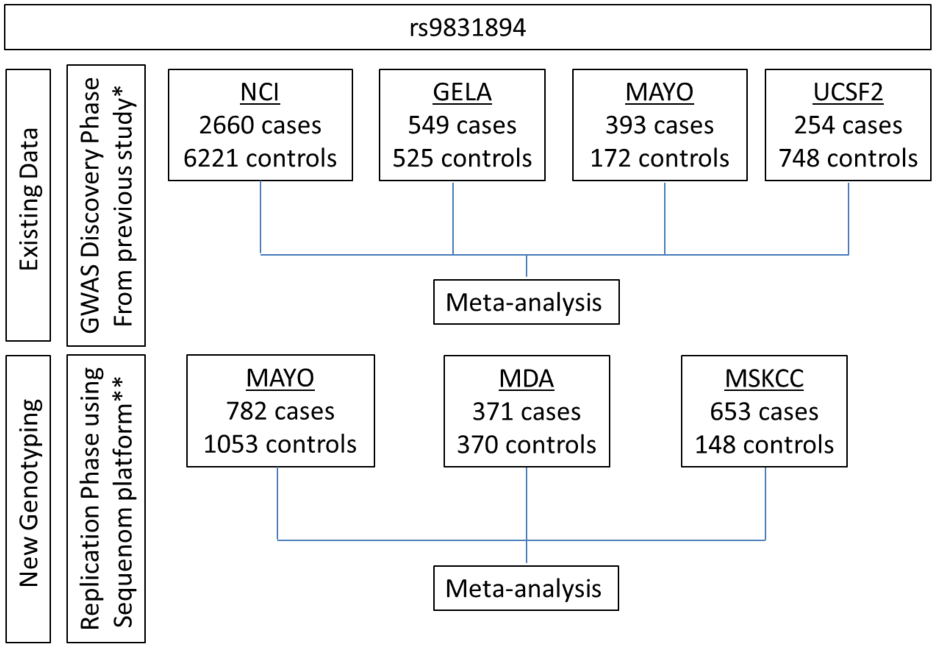
Supplementary Figure 4. Regulatory variants at rs6773363-tagged DLBCL risk locus on 3p24.1. The ATAC-seq data for the 14 blood cell types was downloaded from NCBI GEO under the accession GSE74912. SNP rs2371108 (highlighted), which overlapped a DLBCL super-enhancer, is located within an open chromatin region in nine blood cell types. The alternative T allele at this SNP was predicted to disrupt FOX3, GLI1, KLF3, and ZIC3 binding motif. In addition, SNP rs3806624 (highlighted) is located in an open chromatin region in hematopoietic stem cell. GMP, granulocyte macrophage progenitor; LMPP, lymphoid primed multipotent progenitor; MEP, megakaryocyte erythroid progenitor. See Supplementary Figure 3 legend for additional information.



Supplementary Figure 5. H3K27ac enrichment in DLBCL risk loci across 12 cell types. Except DLBCL patient biopsy, the other 11 cell types are from the 127 reference epigenome, available at https://egg2.wustl.edu/roadmap/data/byFileType/peaks/consolidated/narrowPeak/. Only those whose H3K27ac peaks overlapped at least one of the risk loci were shown. The empirical *p* value was calculated by comparing to a null distribution of randomly sampling minor-allele-frequency-matched variants from the 1000 Genomes Project phase 3 release (Methods). DLBCL risk variants are enriched in DLBCL enhancer as expected, and in enhancers from three immune related cell types as well. E004-E116: reference epigenome identifier.

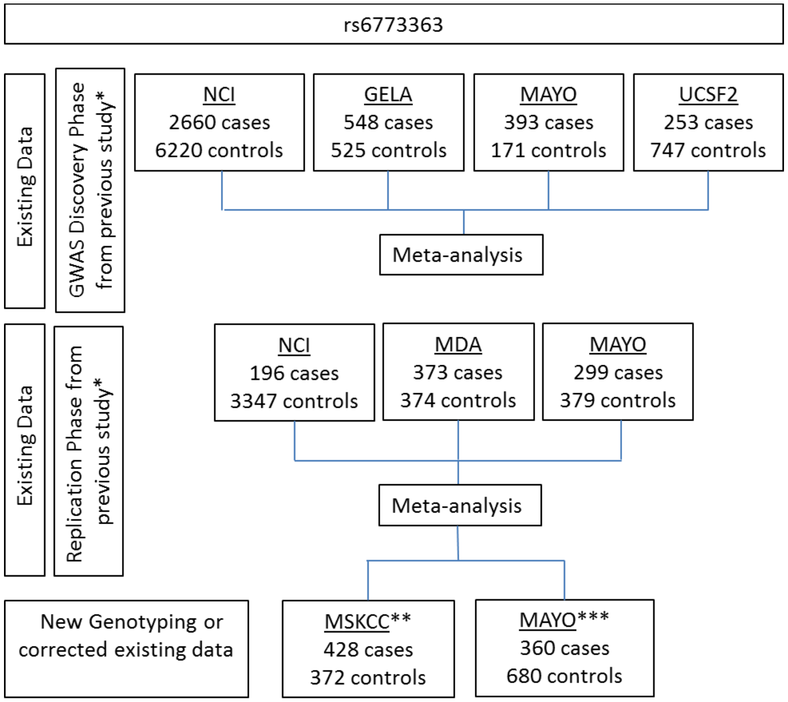


Supplementary Figure 6. Schematic of the study design



*Cerhan,J.R. *et al.* (2014) Genome-wide association study identifies multiple susceptibility loci for diffuse large B cell lymphoma. *Nature genetics*, **46**, 1233–8.

**SNP failed design with Taqman and was unable to be analyzed in the replication stage in our previous GWAS.



*Cerhan,J.R. et al. (2014) Genome-wide association study identifies multiple susceptibility loci for diffuse large B cell lymphoma. Nature genetics, 46, 1233–8.

**SNP miscoded in the original replication stage in our previous GWAS and now corrected

***Additional cases and controls genotyped from Mayo in addition to our previous replication phase

Supplementary References

1. Zhang,J., Vlasevska,S., Wells,V.A., Nataraj,S., Holmes,A.B., Duval,R., Meyer,S.N., Mo,T., Basso,K., Brindle,P.K., *et al.* (2017) The CREBBP Acetyltransferase Is a Haploinsufficient Tumor Suppressor in B-cell Lymphoma. *Cancer discovery*, **7**, 322–337.
2. ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57–74.
3. Qian,J., Wang,Q., Dose,M., Pruett,N., Kieffer-Kwon,K.-R., Resch,W., Liang,G., Tang,Z., Mathé,E., Benner,C., *et al.* (2014) B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. *Cell*, **159**, 1524–37.
4. Abraham,B.J., Cui,K., Tang,Q. and Zhao,K. (2013) Dynamic regulation of epigenomic landscapes during hematopoiesis. *BMC genomics*, **14**, 193.
5. Zhang,J., Dominguez-Sola,D., Hussein,S., Lee,J.-E., Holmes,A.B., Bansal,M., Vlasevska,S., Mo,T., Tang,H., Basso,K., *et al.* (2015) Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nature medicine*, **21**, 1190–8.
6. Chapuy,B., McKeown,M.R., Lin,C.Y., Monti,S., Roemer,M.G.M., Qi,J., Rahl,P.B., Sun,H.H., Yeda,K.T., Doench,J.G., *et al.* (2013) Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer cell*, **24**, 777–90.
7. Ryan,R.J.H., Drier,Y., Whitton,H., Cotton,M.J., Kaur,J., Issner,R., Gillespie,S., Epstein,C.B., Nardi,V., Sohani,A.R., *et al.* (2015) Detection of Enhancer-Associated Rearrangements Reveals Mechanisms of Oncogene Dysregulation in B-cell Lymphoma. *Cancer discovery*, **5**, 1058–71.
8. Abraham,B.J., Hnisz,D., Weintraub,A.S., Kwiatkowski,N., Li,C.H., Li,Z., Weichert-Leahey,N., Rahman,S., Liu,Y., Etchin,J., *et al.* (2017) Small genomic insertions form enhancers that misregulate oncogenes. *Nature Communications*, **8**, 14385.
9. McCabe,M.T., Ott,H.M., Ganji,G., Korenchuk,S., Thompson,C., Van Aller,G.S., Liu,Y., Graves,A.P., Della Pietra,A., Diaz,E., *et al.* (2012) EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature*, **492**, 108–12.
10. Kacmarczyk,T.J., Bourque,C., Zhang,X., Jiang,Y., Houvras,Y., Alonso,A. and Betel,D. (2015) Multiplexing of ChIP-Seq Samples in an Optimized Experimental Condition Has Minimal Impact on Peak Detection. *PloS one*, **10**, e0129350.
11. Care,M.A., Cocco,M., Laye,J.P., Barnes,N., Huang,Y., Wang,M., Barrans,S., Du,M., Jack,A., Westhead,D.R., *et al.* (2014) SPIB and BATF provide alternate determinants of IRF4 occupancy in diffuse large B-cell lymphoma linked to disease heterogeneity. *Nucleic acids research*, **42**, 7591–610.
12. Steube,A., Schenk,T., Tretyakov,A. and Saluz,H.P. (2017) High-intensity UV laser ChIP-seq for the study of protein-DNA interactions in living cells. *Nature communications*, **8**, 1303.
13. Hardee,J., Ouyang,Z., Zhang,Y., Kundaje,A., Lacroute,P. and Snyder,M. (2013) STAT3 targets suggest mechanisms of aggressive tumorigenesis in

diffuse large B-cell lymphoma. *G3 (Bethesda, Md)*, **3**, 2173–85.

14. Ryan,R.J.H., Petrovic,J., Rausch,D.M., Zhou,Y., Lareau,C.A., Kluk,M.J., Christie,A.L., Lee,W.Y., Tarjan,D.R., Guo,B., *et al.* (2017) A B Cell Regulome Links Notch to Downstream Oncogenic Pathways in Small B Cell Lymphomas. *Cell reports*, **21**, 784–797.
15. Koues,O.I., Kowalewski,R.A., Chang,L.-W., Pyfrom,S.C., Schmidt,J.A., Luo,H., Sandoval,L.E., Hughes,T.B., Bednarski,J.J., Cashen,A.F., *et al.* (2015) Enhancer sequence variants and transcription-factor deregulation synergize to construct pathogenic regulatory circuits in B-cell lymphoma. *Immunity*, **42**, 186–98.

Support for individual studies:
ATBC – Intramural Research Program of the National Institutes of Health, NCI, Division of Cancer Epidemiology and Genetics. U.S. Public Health Service contracts (N01-CN-45165, N01-RC-45035, N01-RC-37004, HHSN261201000006C).

BC – Canadian Institutes for Health Research (CIHR). Canadian Cancer Society. Michael Smith Foundation for Health Research.

CPSII – The American Cancer Society funds the creation, maintenance, and updating of the CPSII cohort. The authors thank the CPS-II participants and Study Management Group for their invaluable contributions to this research. The authors would also like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention National Program of Cancer Registries, and cancer registries supported by the National Cancer Institute Surveillance Epidemiology and End Results program.

ELCCS – Leukaemia & Lymphoma Research.

ENGELA – Fondation ARC pour la Recherche sur le Cancer. Fondation de France. French Agency for Food, Environmental and Occupational Health & Safety (ANSES),the French National Cancer Institute (INCa).

EPIC – Coordinated Action (Contract #006438, SP23-CT-2005-006438). HuGeF (Human Genetics Foundation), Torino, Italy.

EPILYMPH – European Commission (grant references QLK4-CT-2000-00422 and FOOD-CT-2006-023103); the Spanish Ministry of Health (grant references CIBERESP, PI11/01810, RCESP C03/09, RTICESP C03/10 and RTIC RD06/0020/0095), the Marató de TV3 Foundation (grant reference 051210), the Agència de Gestiód’AjutsUniversitarisi de Recerca – Generalitat de Catalunya (grant reference 2009SGR1465) who had no role in the data collection, analysis or interpretation of the results; the NIH (contract NO1-CO-12400); the Compagnia di San Paolo—Programma Oncologia; the Federal Office for Radiation Protection grants StSch4261 and StSch4420, the José Carreras Leukemia Foundation grant DJCLS-R12/23, the German Federal Ministry for Education and Research (BMBF-01-EO-1303); the Health Research Board, Ireland and Cancer Research Ireland; Czech Republic supported by MH CZ – DRO (MMCI, 00209805) and RECAMO, CZ.1.05/2.1.00/03.0101; Fondation de France and Association de Recherche Contre le Cancer.

GELA – The French National Cancer Institute (INCa).

HPFS – The HPFS was supported in part by National Institutes of Health grants UM1 CA167552, R01 CA149445 and R01 CA098122. We would like to thank the participants and staff of the Health Professionals Follow-up Study for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data. The study protocol was approved by the institutional review boards of the Brigham and Women’s Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required.

Iowa-Mayo SPORE – National Institutes of Health (CA97274). NCI Specialized Programs of Research Excellence (SPORE) in Human Cancer (P50 CA97274). Molecular Epidemiology of Non-Hodgkin Lymphoma Survival (R01 CA129539). Henry J. Predolin Foundation.

Italian GxE – Italian Ministry for Education, University and Research Research (PRIN 2007 prot. 2007WEJLZB, PRIN 2009 prot. 20092ZELR2); the Italian Association for Cancer Research (AIRC, Investigator Grant 11855). (M.G.E.) - Regional Law N. 7, 2007: “Basic research” (Progetti di ricerca fondamentale o di base) by the Regional Administration of Sardinia (CRP-59812/2012), Fondazione Banco di Sardegna 2010-2012.

Mayo Clinic Case-Control – National Institutes of Health (R01 CA92153). National Center for Advancing Translational Science (UL1 TR000135)

MCCS – The Melbourne Collaborative Cohort Study recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 209057, 251553 and 504711 and by infrastructure provided by Cancer Council Victoria.

MD Anderson – Institutional support to the Center for Translational and Public Health Genomics.

MSKCC – Geoffrey Beene Cancer Research Grant, Lymphoma Foundation (LF5541). Barbara K. Lipman Lymphoma Research Fund (74419). Robert and Kate Niehaus Clinical Cancer Genetics Research Initiative (57470), U01 HG007033. ENCODE, U01 HG007033.

NCI-SEER – Intramural Research Program of the National Cancer Institute, National Institutes of Health, and Public Health Service (N01-PC-65064,N01-PC-67008, N01-PC-67009, N01-PC-67010, N02-PC-71105).

NHS – The NHS was supported in part by National Institutes of Health grants UM1 CA186107, PO1 CA87969, R01 CA49449, R01 CA149445 and R01 CA098122. We would like to thank the participants and staff of the Nurses' Health Study for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data. The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required.

NSW – was supported by grants from the Australian National Health and Medical Research Council (ID990920), the Cancer Council NSW, and the University of Sydney Faculty of Medicine.

NYUWHS - National Cancer Institute (R01 CA098661, P30 CA016087). National Institute of Environmental Health Sciences (ES000260).

PLCO - This research was supported by the Intramural Research Program of the National Cancer Institute and by contracts from the Division of Cancer Prevention, National Cancer Institute, NIH, DHHS.

SCALE – Swedish Cancer Society (2009/659). Stockholm County Council (20110209) and the Strategic Research Program in Epidemiology at Karolinska Institute. Swedish Cancer Society grant (02 6661). Danish Cancer Research Foundation Grant. Lundbeck Foundation Grant (R19-A2364). Danish Cancer Society Grant (DP 08-155). National Institutes of Health (5R01 CA69669-02). Plan Denmark.

UCSF – The UCSF studies were supported by the NCI, National Institutes of Health, CA1046282 and CA154643. The collection of cancer incidence data used in this study was supported by the California Department of Health Services as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology, and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement #1U58 DP000807-01 awarded to the Public Health Institute. The ideas and opinions expressed herein are those of the authors, and endorsement by the State of California, the California Department of Health Services, the National Cancer Institute, or the Centers for Disease Control and Prevention or their contractors and subcontractors is not intended nor should be inferred.

WHI - The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI

investigators can be found at:
http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf

YALE – National Cancer Institute (CA62006).