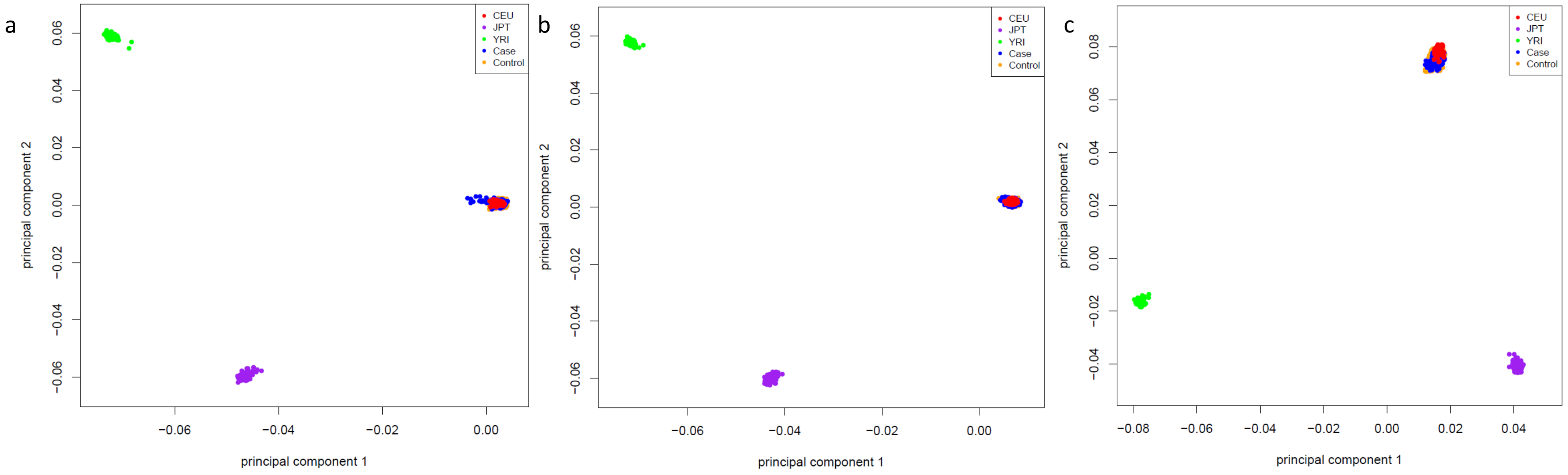
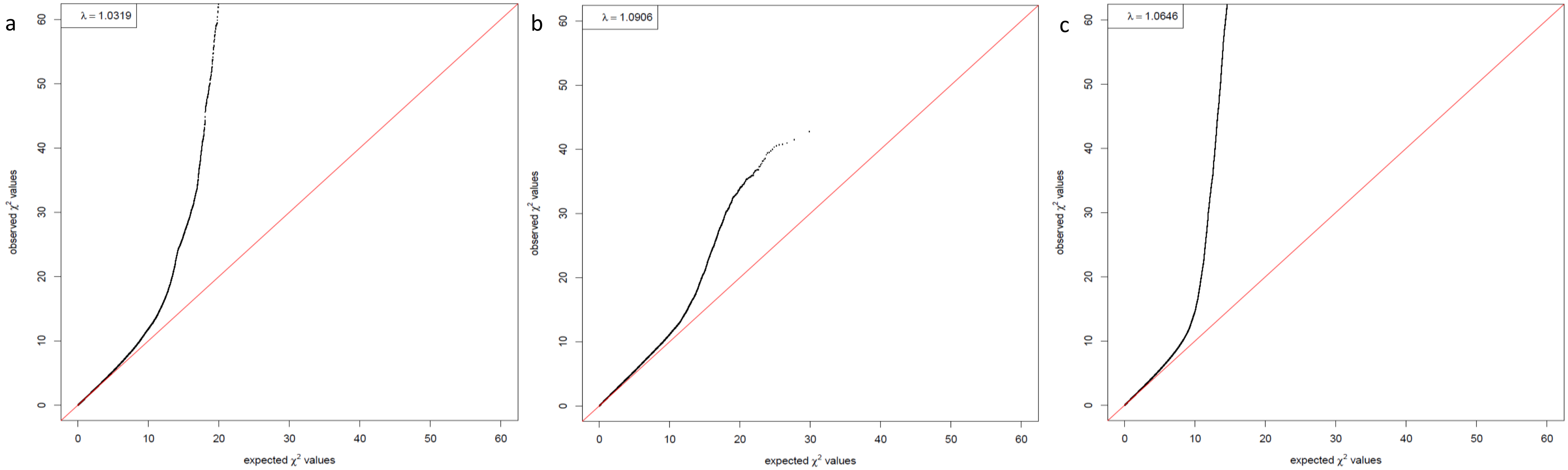
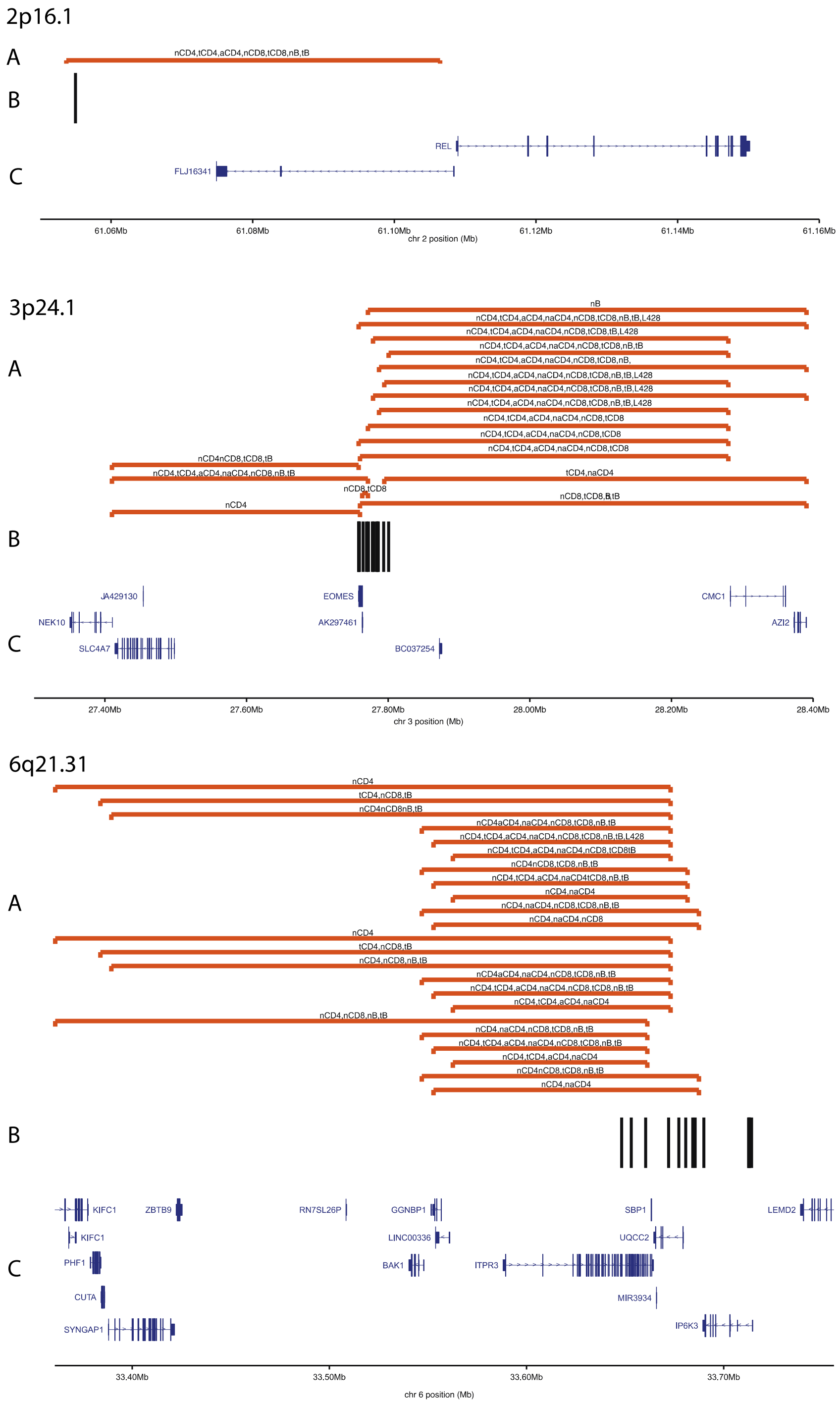
Supplementary figure 1: Identification of individuals of non-European ancestry in cases and controls. (a) UK-GWAS (b) German-GWAS (c) UK-NSHLG-GWAS. The first two principal components of the analysis are plotted. HapMap CEU individuals are plotted in red, JPT individuals are plotted in indigo, YRI are plotted in green. Cases are plotted in blue, controls plotted in orange.

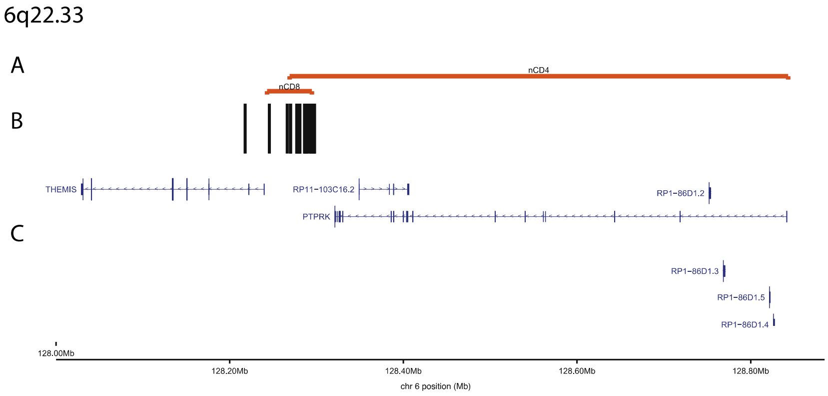


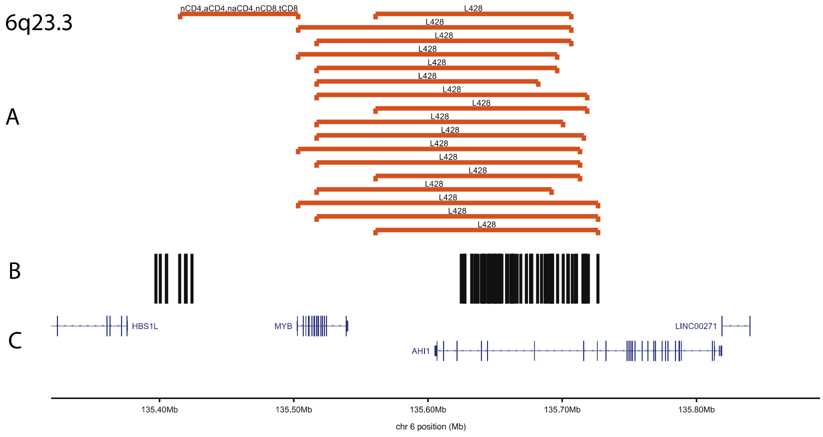
Supplementary Figure 2: Quantile-Quantile (Q-Q) plots of observed and expected χ^2 values of association between SNP genotype and risk of classical Hodgkin lymphoma after imputation. a) UK-GWAS, b) German-GWAS, c) UK-NSHLG-GWAS. The red line represents the null hypothesis of no true association.

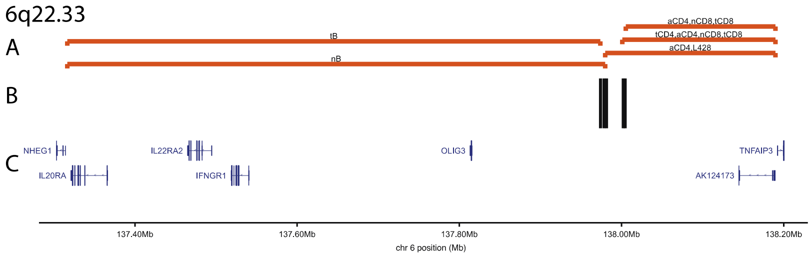


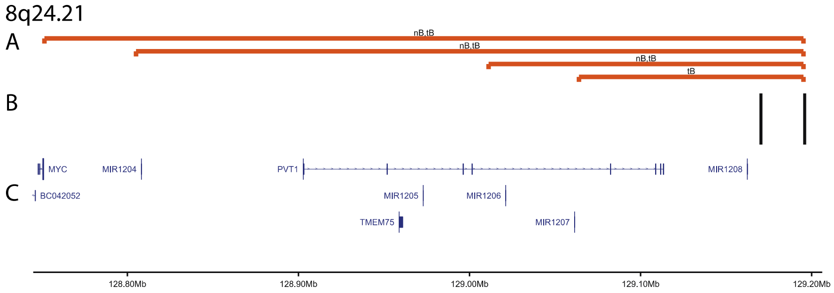
Supplementary Figure 3: Regional plots at which capture Hi-C chromatin contacts overlap HL risk SNPs. For each plot: A, chromatin contact with corresponding cell type; B, HL risk SNPs at each locus (r^2 >0.8); C, gene ID and chromosomal locations.

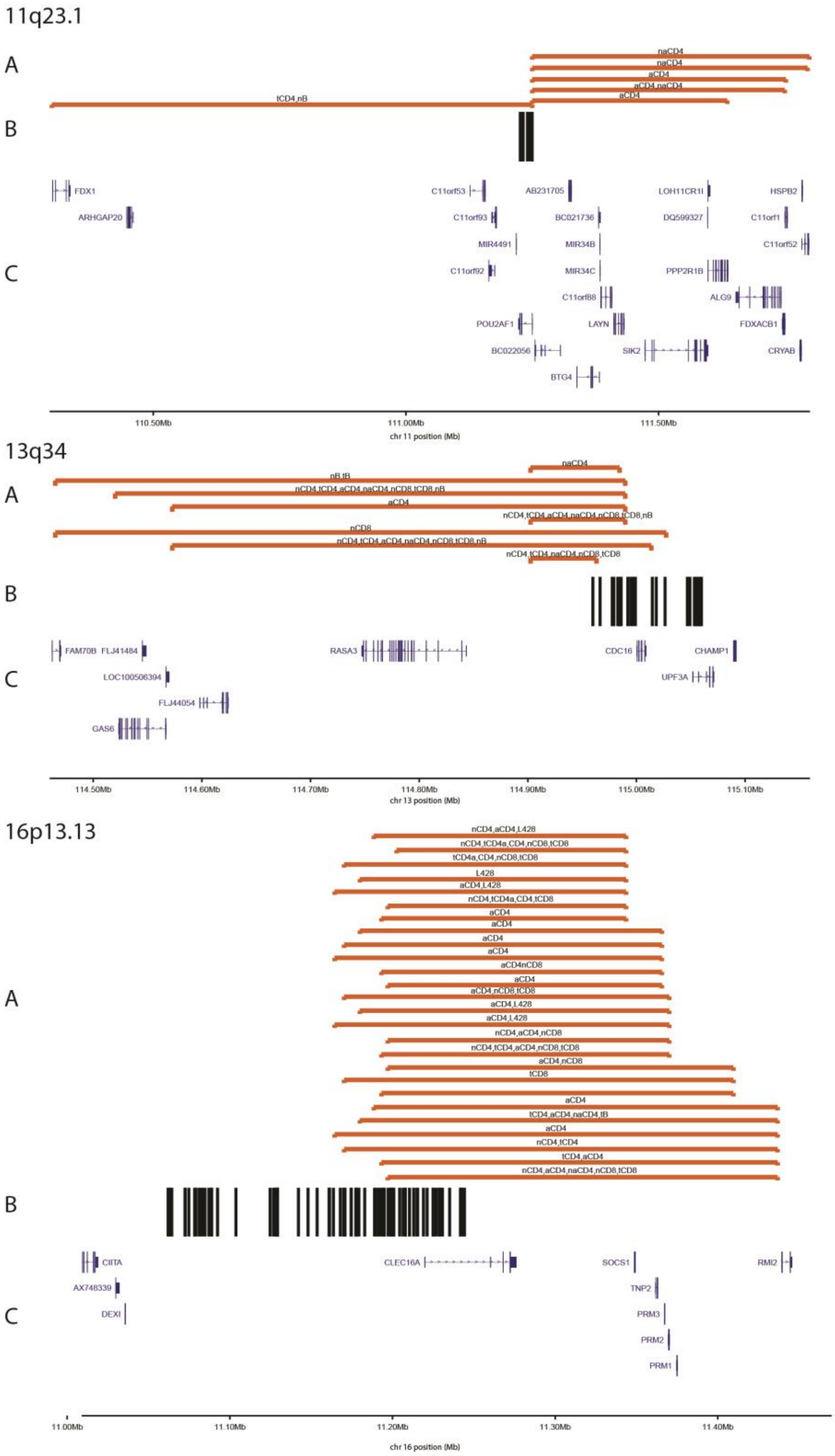


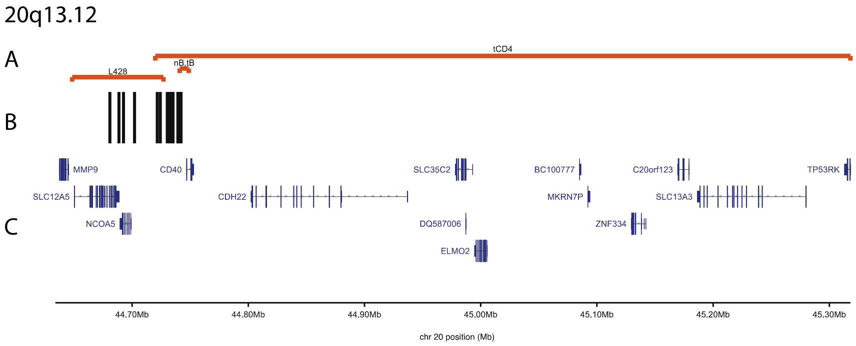

6q22.33
A
nCD4
nCD8
B
C
THEMIS
RP11−103C16.2
PTPRK
RP1−86D1.2
RP1−86D1.3
RP1−86D1.5
RP1−86D1.4
128.00Mb
128.20Mb
128.40Mb
128.60Mb
128.80Mb
chr 6 position (Mb)


6q23.3
A
nCD4,aCD4,naCD4,nCD8,tCD8
L428
B
C
HBS1L
MYB
LINC00271
AHI1
135.40Mb
135.50Mb
135.60Mb
135.70Mb
135.80Mb
chr 6 position (Mb)


6q22.33
A
aCD4,nCD8,tCD8
tCD4,aCD4,nCD8,tCD8
aCD4,L428
tB
nB
B
C
NHEG1
IL22RA2
OLIG3
TNFAIP3
IL20RA
IFNGR1
AK124173
137.40Mb
137.60Mb
137.80Mb
138.00Mb
138.20Mb
chr 6 position (Mb)


8q24.21
A
nB,tB
nB,tB
nB,tB
tB
B
C
MYC
MIR1204
PVT1
MIR1208
BC042052
MIR1205
MIR1206
TMEM75
MIR1207
128.80Mb
128.90Mb
129.00Mb
129.10Mb
129.20Mb


11q23.1
A
B
C
naCD4
naCD4
aCD4
aCD4,naCD4
aCD4
tCD4,nB
FDX1
ARHGAP20
C11orf53
C11orf93
MIR4491
C11orf92
POU2AF1
BC022056
AB231705
BC021736
MIR34B
MIR34C
C11orf88
LAYN
BTG4
SIK2
LOH11CR1I
DQ599327
PPP2R1B
ALG9
FDXACB1
HSPB2
C11orf1
C11orf52
CRYAB
110.50Mb
111.00Mb
111.50Mb
chr 11 position (Mb)
13q34
A
B
C
naCD4
nB,tB
nCD4,tCD4,aCD4,naCD4,nCD8,tCD8,nB
aCD4
nCD4,tCD4,aCD4,naCD4,nCD8,tCD8,nB
nCD8
nCD4,tCD4,aCD4,naCD4,nCD8,tCD8,nB
nCD4,tCD4,naCD4,nCD8,tCD8
FAM70B
FLJ41484
LOC100506394
FLJ44054
GAS6
RASA3
CDC16
CHAMP1
UPF3A
114.50Mb
114.60Mb
114.70Mb
114.80Mb
114.90Mb
115.00Mb
115.10Mb
chr 13 position (Mb)
16p13.13
A
B
C
nCD4,aCD4,L428
nCD4,tCD4a,CD4,nCD8,tCD8
tCD4a,CD4,nCD8,tCD8
L428
aCD4,L428
nCD4,tCD4a,CD4,tCD8
aCD4
aCD4
aCD4
aCD4
aCD4nCD8
aCD4
aCD4,nCD8,tCD8
aCD4,L428
aCD4,L428
nCD4,aCD4,nCD8
nCD4,tCD4,aCD4,nCD8,tCD8
aCD4,nCD8
tCD8
aCD4
tCD4,aCD4,naCD4,tB
aCD4
nCD4,tCD4
tCD4,aCD4
nCD4,aCD4,naCD4,nCD8,tCD8
CIITA
AX748339
DEXI
CLEC16A
SOCS1
TNP2
PRM3
PRM2
PRM1
RMI2
11.00Mb
11.10Mb
11.20Mb
11.30Mb
11.40Mb
chr 16 position (Mb)


20q13.12
A
tCD4
nB,tB
L428
B
C
MMP9
CD40
SLC35C2
BC100777
C20orf123
TP53RK
SLC12A5
CDH22
MKRN7P
SLC13A3
NCOA5
DQ587006
ZNF334
ELMO2
44.70Mb
44.80Mb
44.90Mb
45.00Mb
45.10Mb
45.20Mb
45.30Mb
chr 20 position (Mb)

Supplementary Table 1: Details of each case-control series. Young adult nodular sclerosis Hodgkin lymphoma (NSHL) < 35 years of age at diagnosis. NSHLG, National Study of Hodgkin lymphoma Genetics; cHL, classical Hodgkin lymphoma; MCHL, mixed cellularity Hodgkin lymphoma; EBV, Epstein-Barr virus.

Study	Study centre	Location of enrolment	Cases	Controls
Discovery UK-GWAS[1]	The Institute of Cancer Research	UK	622 cHL cases (58 male, mean age at diagnosis = 24.4 years, 155 NSHL cases) from the Royal Marsden Hospitals National Health Service Trust Family History/ ICR study (2004–2008).	2,930 from the 1958 Birth Cohort. 2,737 from the National Blood Service.
Discovery German-GWAS[2]	The German Cancer Research Centre University of Cologne	Germany	1,001 cHL cases (384 male, mean age at diagnosis = 34.8 years, 402 NSHL cases, 179 MCHL cases, 51 young adult NSHL) from the German Hodgkin study group.	2,092 from the Heinz-Nixdorf Recall Study.
Discovery UK-NSHLG-GWAS[3]	The Institute of Cancer Research	UK	1,717 cHL cases (712 male, mean age at diagnosis = 41.5 years, 720 NSHL cases, 637 MCHL cases, 261 young adult NSHL) from the National Study of Hodgkin lymphoma Genetics (2010-2013).	2,976 from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium 4,446 from the Breast Cancer Association Consortium.
UK-Biobank[4]	UK-Biobank	UK	432 HL cases from the UK Biobank.	337,208 individuals from the UK Biobank.
USC-IARC-UC-GWAS[5,6]	University of Southern California International Agency for Research on Cancer University of Chicago	Europe and USA	1,816 HL cases (921 male, 1233 NSHL cases, 320 MCHL cases, 287 EBV positive cases, 776 EBV negative cases).	7,879 controls.

Supplementary Table 2: Details of the quality control filters applied to each genome-wide association study. Samples were excluded due to call rate (< 95% or failed genotyping), ethnicity (principle components analysis or other samples reported to be not of white, European descent), relatedness (any individuals found to be duplicated or related within or between data sets through identity by state) or sex discrepancy. NSHLG, National Study of Hodgkin lymphoma Genetics.

	UK-GWAS		German-GWAS		UK-NSHLG-GWAS	
	Cases	Controls	Cases	Controls	Cases	Controls
Pre-quality control	622	5,677	1,001	2,092	1,717	7,519
Sex discrepancy	3	25	11	0	14	0
Call rate fail	12	248	9	0	3	12
Heterozygosity rate	NA	NA	27	NA	25	43
Related Individuals	2	71	19	0	14	66
None-European Ancestry	30	109	86	0	59	135
Post-quality control[†]	589	5,197	876	2,092	1,612	7,263

[†] filters for quality control were performed simultaneously so numbers for each criteria may not sum to total removed.

Supplementary Table 3: Histology of the cases in each genome-wide association study (GWAS) and replication. NSHLG, National Study of Hodgkin lymphoma Genetics.

	UK-GWAS	German-GWAS	UK-NSHLG-GWAS
Classical Hodgkin lymphoma	589	876	1,612
Nodular sclerosis Hodgkin lymphoma	155 (26%)	402 (46%)	720 (45%)
Mixed cellularity Hodgkin lymphoma	12 (2%)	179 (20%)	637 (40%)
Lymphocyte-rich Hodgkin lymphoma	5 (1%)	0 (0%)	7 (<1%)
Lymphocyte-depleted Hodgkin lymphoma	8 (1%)	0 (0%)	0 (0%)
Unspecified	421 (71%)	295 (34%)	248 (15%)

Supplementary Table 4: Details of the quality control filters applied to each genome-wide association (GWAS) study. Genotyped single nucleotide polymorphisms (SNPs) with a call rate < 95% were excluded as were those with a minor allele frequency (MAF) < 0.01 or displaying significant deviation from Hardy-Weinberg equilibrium (HWE) (i.e. $P < {}^{10-5}$). 1958BC, 1958 birth cohort; UKBS, UK blood service; NSHLG, National Study of Hodgkin lymphoma Genetics.

	UK-GWAS			German-GWAS	UK-NSHLG-GWAS
	Cases	Controls			
		1958BC	UKBS		
Genotyping Platform	Illumina 660W-Quad BeadChips	Illumina Human1.2M-Duo Custom_v1 Array BeadChips		Illumina Human OmniExpress-12 v1.0	Infinium OncoArray-500K BeadChip
Pre-quality control	640,663	1,115,428	1,115,428	733,202	464,624
Call rate fail	99,296	18,418	20,224	2,083	11
HWE fail	867	11,662	14,374	1,104	14
MAF < 0.01	24,177	179,500	178,713	78,364	73,351
Other*	n/a	189,910	183,751	n/a	n/a
Post-quality control[†]	516,323	899,696	900,580	649,788	382,935

[*] info score and plate association (see www.wtccc.org.uk for details)

[†] filters for QC were performed simultaneously so numbers for each criteria may not sum to total removed.

Supplementary Table 5: Relationship between new Hodgkin lymphoma risk single nucleotide polymorphisms (SNP) and histology of Hodgkin lymphoma cases. *P*-value calculated from logistic regression in each individual dataset followed by a standard meta-analysis. NSHL, nodular sclerosis Hodgkin lymphoma; MCHL, mixed cellularity Hodgkin lymphoma; bp, base pair; OR, odds ratio; CI, confidence interval; I^2 proportion of the total variation due to heterogeneity.

						Meta-analysis		
	Locus	Nearest genes†	Risk allele	Position (hg19, bp)	*P* value	OR (95% CI)	I^2 (%)	P_{het}
	6p21.31, rs649775	*UQCC2-IP6K3*	A	33684313				
NSHL					7.40×10^{-4}	1.24 (1.06-1.45)	72	0.03
MCHL					0.01	1.25 (1.05-1.52)	0	0.75
Case only analysis					0.47	0.92 (0.74-1.15)	0	0.48
	6q23.3, rs1002658	*OLIG3-TNFAIP3*	T	137981584				
NSHL					4.04×10^{-4}	1.22 (1.09-1.36)	0	0.41
MCHL					0.12	1.11 (0.97-1.27)	0	0.78
Case only analysis					0.52	1.06 (0.90-1.24)	0	0.83
	11q23.1, rs7111520	*POU2AF1*	A	111249611				
cHL								
NSHL					2.68×10^{-3}	1.14 (1.05-1.25)	0	0.68
MCHL					3.74×10^{-3}	1.17 (1.05-1.31)	20	0.26
Case only analysis					0.67	0.97 (0.84-1.12)	0	0.62
	16p11.2, rs6565176	*MAPK3-CORO1A*	T	30174926				
NSHL					0.02	1.10 (1.01-1.19)	40	0.19
MCHL					0.03	1.12 (1.02-1.24)	0	0.37
Case only analysis					0.94	1.00 (0.88-1.15)	50	0.16
	20q13.12, rs2425752	*NCOA5-CD40*	T	44702120				
NSHL					2.27×10^{-3}	1.16 (1.05-1.27)	0	0.98
MCHL					0.71	0.98 (0.87-1.10)	49	0.16
Case only analysis					0.04	1.16 (1.00-1.33)	38	0.21

Supplementary Table 6: Relationship between new Hodgkin lymphoma risk single nucleotide polymorphisms (SNP) and sex of Hodgkin lymphoma cases, young adult nodular sclerosis Hodgkin lymphoma (16-35). *P*-value calculated from logistic regression in each individual dataset followed by a standard meta-analysis where data from multiple studies.

[1] UK-GWAS (men=58, women=531), German-GWAS (men=348, women=528), UK-NSHLG-GWAS (men=712, women=900).

[2] German-GWAS (young adult NSHL=151, other NSHL=251), UK-NSHLG-GWAS (young adult NSHL=261, other NSHL=459).

		Sex[1]	Age[2]
Locus	SNP	*P* value	
6p21.31	rs649775	0.46	0.81
6q23.3	rs1002658	0.43	0.16
11q23.1	rs7111520	0.94	0.41
16p11.2	rs6565176	0.54	0.34
20q13.12	rs2425752	0.01	0.98

Supplementary Table 7: Results of enrichment of HL risk loci at H3K4Me3 peaks in 141 haematological cell types.

Cell	***P* value**
CD4+ T-cell	0.00296
CD3- CD4+ CD8+ thymocyte	0.00574
Germinal center B-cell	0.0063
CD3- CD4+ CD8+ thymocyte	0.00874
CD3+ CD4+ CD8+ thymocyte	0.00914
CD3+ CD4+ CD8+ thymocyte	0.00926
Effector memory CD8+ T-cell	0.01504
Cytotoxic CD56dim NK cell	0.01676
CD4+ T-cell	0.02252
CD8+ T-cell	0.02712
CD8+ T-cell	0.0302
Central memory CD4+ T-cell	0.03234
CD4+ thymocyte	0.04428
CD4+ T-cell	0.05268
Class switched memory B-cell	0.05468
CD34- CD41+ CD42+ megakaryocyte	0.05748
CD4+ T-cell	0.06064
CD3+ CD4+ CD8+ thymocyte	0.08102
Endothelial cell of umbilical vein (proliferating)	0.09658
Germinal center B-cell	0.11258
Naïve B-cell	0.12136
Naïve B-cell	0.12972
CD4+ T-cell	0.14058
CD38- Naïve B-cell	0.16362
CD14+ CD16- classical monocyte	0.17008
Endothelial cell of umbilical vein (resting)	0.1735
Alternatively activated macrophage	0.18746
Effector memory CD8+ T-cell	0.19482
L-428	0.2037
Inflammatory macrophage	0.20728
Plasma cell	0.21626
Neutrophilic myelocyte	0.23564
Cytotoxic CD56dim NK cell	0.25434
Cytotoxic CD56dim NK cell	0.3167
Band form neutrophil	0.3206
CD8+ T-cell	0.3297
CD38- Naïve B-cell	0.35362
CD38- Naïve B-cell	0.40246
CD4+ T-cell	0.41788
Effector memory CD4+ T-cell	0.4564
Neutrophilic metamyelocyte	0.46298
Cytotoxic CD56dim NK cell	0.46912
Naïve B-cell	0.4774
CD38- Naïve B-cell	0.47786
Effector memory CD8+ T-cell (terminally differentiated)	0.48562
Macrophage	0.51654
Germinal center B-cell	0.51828
Naïve B-cell	0.52458
Naïve B-cell	0.5486
Cytotoxic CD56dim NK cell	0.54922
Endothelial cell of umbilical vein (resting)	0.55432
Effector memory CD8+ T-cell	0.55596
CD8+ T-cell	0.56432
Mature neutrophil	0.56764
CD8+ thymocyte	0.57826
Central memory CD4+ T-cell	0.58364
Mature eosinophil	0.59558
CD14+ CD16- classical monocyte	0.61336

Band form neutrophil	0.62058
Cytotoxic CD56dim NK cell	0.62402
Germinal center B-cell	0.63534
CD14+ CD16- classical monocyte	0.63922
CD4+ T-cell	0.64242
Naïve B-cell	0.648
CD8+ T-cell	0.6591
Naïve B-cell	0.66082
CD38- Naïve B-cell	0.66482
CD34- CD41+ CD42+ megakaryocyte	0.66998
Alternatively activated macrophage	0.67934
Plasma cell	0.68046
Erythroblast	0.69146
CD34- CD41+ CD42+ megakaryocyte	0.72396
Segmented neutrophil	0.72456
Osteoclast	0.73698
Unswitched memory B-cell	0.74876
Plasma cell	0.74912
Neutrophilic metamyelocyte	0.75756
Inflammatory macrophage	0.75938
Segmented neutrophil	0.77034
Class switched memory B-cell	0.77394
CD14+ CD16- classical monocyte	0.79614
Naïve B-cell	0.80356
Mature neutrophil	0.80508
CD4+ T-cell	0.80566
CD14+ CD16- classical monocyte	0.8173
Class switched memory B-cell	0.81788
Endothelial cell of umbilical vein (proliferating)	0.82004
Inflammatory macrophage	0.8217
Macrophage	0.82382
Band form neutrophil	0.82586
CD38- Naïve B-cell	0.83326
Alternatively activated macrophage	0.83948
Neutrophilic metamyelocyte	0.84056
CD14+ CD16- classical monocyte	0.85394
Macrophage	0.85484
Mature neutrophil	0.85774
Macrophage	0.8819
Mature neutrophil	0.88676
CD4+ T-cell	0.88832
Osteoclast	0.8898
CD14+ CD16- classical monocyte	0.89142
CD14+ CD16- classical monocyte	0.89646
Conventional dendritic cell	0.89676
Mature neutrophil	0.9019
CD14+ CD16- classical monocyte	0.9252
Macrophage	0.92812
Segmented neutrophil	0.9298
Neutrophilic myelocyte	0.92984
Inflammatory macrophage	0.93216
Mature neutrophil	0.93582
Neutrophilic myelocyte	0.9383
Inflammatory macrophage	0.9395
Neutrophilic myelocyte	0.94164
Inflammatory macrophage	0.9458
Segmented neutrophil	0.94628
Mature neutrophil	0.94892
CD14+ CD16- classical monocyte	0.95094
CD38- Naïve B-cell	0.95434
Mature neutrophil	0.95936
Inflammatory macrophage	0.96148
Mature eosinophil	0.96326

Erythroblast	0.96406
Mature neutrophil	0.9703
Mature neutrophil	0.9707
Mature neutrophil	0.97312
Mesenchymal stem cell of the bone marrow	0.9741
Macrophage	0.97912
Macrophage	0.98136
Macrophage	0.98146
Mesenchymal stem cell of the bone marrow	0.982
Alternatively activated macrophage	0.98484
Mature neutrophil	0.98648
Neutrophilic metamyelocyte	0.98718
Alternatively activated macrophage	0.9875
Band form neutrophil	0.98762
Alternatively activated macrophage	0.9898
Alternatively activated macrophage	0.9902
Adult endothelial progenitor cell	0.99182
Adult endothelial progenitor cell	0.99182
Mature neutrophil	0.99432
Inflammatory macrophage	0.99462

Supplementary Table 8: Results of enrichment of HL risk loci at H3K27Ac peaks in 126 haematological cell types.

Cell	*P* Value
CD8+ T-cell	0.0003
CD3+ CD4+ CD8+ thymocyte	0.00056
CD4+ thymocyte	0.00266
L-428	0.0079
CD4+ T-cell	0.00922
Alternatively activated macrophage	0.01554
CD3- CD4+ CD8+ thymocyte	0.0243
CD4+ T-cell	0.04416
CD38- naïve B-cell	0.04574
CD4+ T-cell	0.0622
CD8+ T-cell	0.06528
CD38- naïve B-cell	0.0689
Endothelial cell umbilical vein	0.07034
CD4+ T-cell	0.07108
Plasma cell	0.08442
CD8+ T-cell	0.08568
CD4+ T-cell	0.0892
CD8+ T-cell	0.09292
CD8+ T-cell	0.0953
Cytotoxic CD56dim NK cell	0.11062
Germinal centre B-cell	0.11116
Germinal centre B-cell	0.11442
Erythoblast	0.12128
Class switched memory B-cell	0.1214
Cytotoxic CD56dim NK cell	0.13596
CD4+ T-cell	0.14956
CD38- naïve B-cell	0.1769
Cytotoxic CD56dim NK cell	0.18084
Inflammatory macrophage	0.1894
CD4+ T-cell	0.19676
Osteoclast	0.19678
Naïve B-cell	0.2124
CD3+ CD4+ CD8+ thymocyte	0.21358
Mature neutrophil	0.24172
Erythoblast	0.28154
Mature neutrophil	0.2857
Naïve B-cell	0.29322
Unswitched memory B-cell	0.30118
Effector memory CD8+ T-cell	0.30154
Naïve B-cell	0.32602
CD4+ T-cell	0.36336
Macrophage	0.36516
CD8+ T-cell	0.37626
Osteoclast	0.3914
Mature neutrophil	0.41226
Inflammatory macrophage	0.4135
Segmented neutrophil	0.4282
Segmented neutrophil	0.43532
Mature neutrophil	0.44386
Germinal centre B-cell	0.44566
Naïve B-cell	0.448
Mature neutrophil	0.4574
Mature eosinophil	0.47724
Effector memory CD8+ T-cell	0.48382
CD34- CD41+ CD42+ megakaryocyte	0.48402
Cytotoxic CD56dim NK cell	0.49758
Class switched memory B-cell	0.51014
CD14+ CD16- classical monocyte	0.51158

Naïve B-cell	0.52138
Plasma cell	0.5459
Endothelial cell of umbilical vein (proliferating)	0.55018
Class switched memory B-cell	0.55168
Alternatively activated macrophage	0.55928
CD38- naïve B-cell	0.57236
Alternatively activated macrophage	0.59688
Adult endothelial progenitor cell	0.60512
Naïve B-cell	0.62676
Central memory CD4+ T-cell	0.6336
Macrophage	0.63856
Naïve B-cell	0.6492
CD4+ T-cell	0.65258
Plasma cell	0.65838
Naïve B-cell	0.72982
Endothelial cell of umbilical vein (proliferating)	0.75132
CD14+ CD16- classical monocyte	0.7897
Mature neutrophil	0.79488
Mesenchymal stem cell	0.79578
Mature neutrophil	0.80014
CD34- CD41+ CD42+ megakaryocyte	0.80018
Mature neutrophil	0.81364
Mature neutrophil	0.81436
CD14+ CD16- classical monocyte	0.83204
Mature neutrophil	0.84938
Inflammatory macrophage	0.85612
Neutrophilic metamyelocyte	0.86222
Alternatively activated macrophage	0.86706
Mesenchymal stem cell	0.86916
CD14+ CD16- classical monocyte	0.89352
Alternatively activated macrophage	0.90648
Mature eosinophil	0.90884
Alternatively activated macrophage	0.9126
CD14+ CD16- classical monocyte	0.91312
Macrophage	0.9144
Mature neutrophil	0.91984
Band form neutrophil	0.92566
Mature neutrophil	0.93216
Neutrophilic myelocyte	0.9384
Macrophage	0.94828
Conventional dendritic cell	0.94972
CD14+ CD16- classical monocyte	0.95272
Macrophage	0.95618
Adult endothelial progenitor cell	0.95984
Neutrophilic myelocyte	0.96408
Alternatively activated macrophage	0.96788
CD14+ CD16- classical monocyte	0.96966
CD14+ CD16- classical monocyte	0.97622
Inflammatory macrophage	0.97688
Inflammatory macrophage	0.97886
CD14+ CD16- classical monocyte	0.97958
Neutrophilic metamyelocyte	0.98252
Inflammatory macrophage	0.98848
Segmented neutrophil	0.9901
CD14+ CD16- classical monocyte	0.99178
Neutrophilic metamyelocyte	0.992
Inflammatory macrophage	0.99402
Inflammatory macrophage	0.99432
CD14+ CD16- classical monocyte	0.99548
Band form neutrophil	0.99562
Macrophage	0.99578
Band form neutrophil	0.99636
Neutrophilic myelocyte	0.99762

Macrophage	0.9982
Macrophage	0.99822
CD14+ CD16- classical monocyte	0.99854
CD14+ CD16- classical monocyte	0.9986
CD14+ CD16- classical monocyte	0.99898

Supplementary Table 9: Results of expression quantitative trait loci analysis at HL risk loci in lymphoblastoid B-cells. Shown are results that are statistically significant following correction for multiple testing ($P < 0.05$/number by of transcripts (RNA-seq) or number of probes (microarray) at each given locus. Chr, chromosome; eqtl_b expression quantitive trait locus beta; eqtl_se expression quantitative trait locus standard error; gwas_b, genome-wide association study beta; gwas_se, genome-wide association study standard error.

Dataset	SNP	Chr	Position	Effect Allele	Other Allele	Probe	Probe Chr	Probe position	Gene	eqtl_b	eqtl_se	eqtl_p	gwas_b	gwas_se	gwas_p
MuTHER	rs942637	6	33653111	C	T	ILMN_1815500	6	33588142	*ITPR3*	0.084	0.025	7.458E-04	-0.206	0.046	6.34E-06
GTex	rs9402681	6	135397301	T	G	ENSG00000118514.9	6	135238528	*ALDH8A1*	-0.663	0.107	1.662E-08	-0.119	0.029	3.62E-05
MuTHER	rs1331308	6	135405122	A	C	ILMN_1699258	6	135238528	*ALDH8A1*	-0.027	0.005	1.187E-07	-0.110	0.029	1.26E-04
MuTHER	rs2614276	6	135681704	C	T	ILMN_1791006	6	135604670	*AHI1*	0.278	0.019	2.680E-50	0.162	0.029	3.42E-08
GTex	rs10872430	6	135704092	C	T	ENSG00000135541.16	6	135604670	*AHI1*	0.634	0.103	1.788E-08	0.162	0.029	2.00E-08
MuTHER	rs3824662	10	8104208	A	C	ILMN_2406656	10	8095567	*GATA3*	0.102	0.021	6.577E-07	-0.241	0.039	5.38E-10
GTex	rs4938509	11	111246611	T	G	ENSG00000196167.5	11	111164114	*COLCA1*	0.627	0.136	1.338E-05	0.152	0.031	6.70E-07
GTex	rs7111520	11	111249611	A	G	ENSG00000214290.3	11	111169565	*COLCA2*	0.561	0.112	2.895E-06	0.148	0.031	1.28E-06
GTex	rs17067943	13	115054448	T	A	ENSG00000169062.10	13	115047059	*UPF3A*	-0.917	0.128	1.868E-10	0.172	0.057	2.57E-03
MuTHER	rs17067943	13	115054448	T	A	ILMN_1671742	13	115047059	*UPF3A*	-0.137	0.034	7.079E-05	0.172	0.057	2.57E-03
MuTHER	rs7337695	13	115056740	G	C	ILMN_2339796	13	115000362	*CDC16*	0.083	0.020	2.147E-05	0.168	0.057	3.36E-03
MuTHER	rs7184491	16	11170761	C	T	ILMN_1774733	16	11348262	*SOCS1*	0.097	0.029	8.333E-04	0.122	0.035	4.78E-04
GTex	rs35300161	16	11198932	T	C	ENSG00000185338.4	16	11348262	*SOCS1*	0.275	0.084	1.471E-03	0.132	0.036	2.24E-04
GTex	rs2887799	16	30163833	A	G	ENSG00000102882.7	16	30125426	*MAPK3*	0.505	0.076	1.729E-09	-0.115	0.029	8.59E-05
GTex	rs12102320	16	30178516	C	A	ENSG00000183336.7	16	29464914	*BOLA2*	0.820	0.116	2.814E-10	-0.122	0.029	2.09E-05
MuTHER	rs11644392	16	30328317	C	T	ILMN_1667260	16	30116131	*MAPK3*	-0.101	0.013	3.426E-15	0.133	0.029	6.83E-06
MuTHER	rs6032662	20	44734310	C	T	ILMN_1672076	20	44313292	*WFDC10B*	-0.018	0.005	9.069E-04	-0.136	0.033	3.09E-05

Supplementary Table 10: Results of expression quantitative trait loci analysis at HL risk loci in CD4+ and CD8+ T-cells. Shown are results that are statistically significant following correction for multiple testing ($P < 0.05$/number of probes (microarray) at each given locus. Chr, chromosome; gwas_b, genome-wide association study beta; gwas_se, genome-wide association study standard error.

T-cell	SNP	Chr	Position	Effect Allele	Other allele	ILMNid_HT12v4	Probe Chr	Probe centre position	Gene	eQTL Z score	eQTL_*p*	gwas_b	gwas_se	gwas_p
CD4+	rs6570001	6	135651721	C	G	ILMN_1791006	6	135605306	AHI1	10.478	1.09E-25	0.165	0.030	2.41E-08
CD4+	rs9562094	13	114992075	G	A	ILMN_1727389	13	115038008	CDC16	5.955	2.60E-09	0.131	0.053	1.32E-02
CD8+	rs6773363	3	27793632	C	T	ILMN_1760509	3	27757954	EOMES	4.821	1.43E-06	0.212	0.029	2.86E-13
CD8+	rs2246852	6	135691792	G	A	ILMN_1791006	6	135605306	AHI1	-8.146	3.74E-16	-0.162	0.029	3.67E-08
CD8+	rs11147321	13	114966440	G	A	ILMN_2339796	13	115038019	CDC16	5.596	2.20E-08	0.145	0.052	5.53E-03

Supplementary Table 11: Genetic correlation of Hodgkin lymphoma (cHL), nodular sclerosis Hodgkin lymphoma (NSHL) and mixed cellularity Hodgkin lymphoma (MCHL) with infectious diseases. se, standard error; r_g, genetic correlation.

	HL			NSHL			MCHL		
Infection[7]	r_g	se	*P* value	r_g	se	*P* value	r_g	se	*P* value
Chickenpox	-0.03	0.14	0.83	-0.10	0.15	0.48	0.08	0.18	0.67
Shingles	-0.24	0.17	0.16	-0.03	0.17	0.89	-0.23	0.59	0.69
Cold sores	0.09	0.21	0.30	-0.04	0.21	0.80	0.05	0.40	0.70
Infectious mononucleosis	0.34	0.22	0.12	0.20	0.21	0.34	0.26	0.42	0.54
Mumps	0.24	0.14	0.08	0.13	0.14	0.35	0.17	0.22	0.42
Hepatitis A	0.49	0.76	0.52	0.12	0.51	0.82	-0.03	1.62	0.99
Hepatitis B	-0.36	0.31	0.23	-0.65	0.34	0.07	0.01	0.43	0.98
Plantar warts	-0.03	0.20	0.87	0	0.22	0.99	-0.07	0.30	0.82
Positive tuberculosis test	0.06	0.20	0.78	-0.08	0.20	0.68	0.36	0.35	0.31
Streptococcal throat infections	0.10	0.13	0.43	0.18	0.13	0.18	-0.10	0.20	0.62
Scarlett fever	-0.19	0.18	0.29	-0.14	0.19	0.46	0.05	0.43	0.91
Pneumonia	0.15	0.12	0.19	0.15	0.11	0.16	0.07	0.29	0.82
Bacterial meningitis	-0.16	0.34	0.63	-0.18	0.31	0.56	0	0.49	0.99
Yeast infection	0.10	0.11	0.35	0.14	0.11	0.22	0	0.18	0.98
Urinary tract infection	-0.21	0.11	0.06	-0.18	0.12	0.12	-0.20	0.20	0.32
Tonsillectomy	0.05	0.08	0.57	0.09	0.09	0.32	0.02	0.10	0.86
Childhood ear infection	0	0.08	0.95	-0.01	0.09	0.88	0.04	0.11	0.69
Myringotomy	0.12	0.13	0.36	0.13	0.14	0.35	0.21	0.21	0.32
Measles	0.16	0.25	0.53	-0.03	0.27	0.92	0.53	0.46	0.25
Rheumatic fever	-0.03	0.22	0.82	0	0.29	0.78	0.07	0.31	0.80
Common cold	-0.10	0.14	0.46	0.02	0.15	0.89	-0.30	0.21	0.15
Rubella	-0.22	0.16	0.15	-0.16	0.17	0.33	-0.16	0.28	0.56
Chronic sinus infection	0.27	0.19	0.15	0.15	0.19	0.43	-0.31	0.25	0.22

Supplementary Table 12: Assessment of the causal relationship between infection and Hodgkin lymphoma (HL), nodular sclerosis Hodgkin lymphoma (NSHL) and mixed cellularity Hodgkin lymphoma (MCHL) using Mendelian randomisation. Power calculated as per Brion *et al.* assuming an α=0.05 and an odds ratio of 1.5

	HL				NSHL				MCHL			
Infection	**Estimate**	**Standard error**	***P* value**	**Power**	**Estimate**	**Standard error**	***P* value**	**Power**	**Estimate**	**Standard error**	***P* value**	**Power**
Tonsillectomy	0.11	0.16	0.49	0.57	-0.09	0.20	0.64	0.47	0.06	0.21	0.76	0.42
Childhood ear infection	-0.16	0.16	0.37	0.35	-0.65	0.54	0.21	0.26	-0.49	0.24	0.05	0.24
Mumps	0.24	0.11	0.04	0.43	0.28	0.17	0.10	0.35	0.36	0.21	0.08	0.08
Yeast infection	-0.17	0.40	0.68	0.10	-0.56	0.59	0.34	0.09	-0.02	0.73	0.98	0.08

REFERENCES

1. Enciso-Mora V, Broderick P, Ma Y, et al. A genome-wide association study of Hodgkin's lymphoma identifies new susceptibility loci at 2p16.1 (REL), 8q24.21 and 10p14 (GATA3). *Nat Genet*. 2010;42(12):1126-1130.
2. Frampton M, da Silva Filho MI, Broderick P, et al. Variation at 3p24.1 and 6q23.3 influences the risk of Hodgkin's lymphoma. *Nat Commun*. 2013;4:2549.
3. Sud A, Thomsen H, Law PJ, et al. Genome-wide association study of classical Hodgkin lymphoma identifies key regulators of disease susceptibility. *Nature Communications*. 2017;8(1):1892.
4. DeBoever C, Tanigawa Y, McInnes G, et al. Medical relevance of protein-truncating variants across 337,208 individuals in the UK Biobank study. *bioRxiv*. 2017.
5. Cozen W, Timofeeva MN, Li D, et al. A meta-analysis of Hodgkin lymphoma reveals 19p13.3 TCF3 as a novel susceptibility locus. *Nature Communications*. 2014;5:3856.
6. Khankhanian P, Cozen W, Himmelstein DS, et al. Meta-analysis of genome-wide association studies reveals genetic overlap between Hodgkin lymphoma and multiple sclerosis. *Int J Epidemiol*. 2016;45(3):728-740.
7. Tian C, Hromatka BS, Kiefer AK, et al. Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. *Nature Communications*. 2017;8(1):599.