REPORT

Genome-wide Association Study Identifies Five Susceptibility Loci for Follicular Lymphoma outside the HLA Region

Christine F. Skibola,[1,2,84,*] Sonja I. Berndt,[3,84] Joseph Vijai,[4,84] Lucia Conde,[1,2,84] Zhaoming Wang,[5,84] Meredith Yeager,[5,84] Paul I.W. de Bakker,[6,7] Brenda M. Birmann,[8] Claire M. Vajdic,[9] Jia-Nee Foo,[10] Paige M. Bracci,[11] Roel C.H. Vermeulen,[7,12] Susan L. Slager,[13] Silvia de Sanjose,[14,15] Sophia S. Wang,[16] Martha S. Linet,[3] Gilles Salles,[17,18,19] Qing Lan,[3] Gianluca Severi,[20,21,22] Henrik Hjalgrim,[23] Tracy Lightfoot,[24] Mads Melbye,[23,25] Jian Gu,[26] Hervé Ghesquières,[19,27] Brian K. Link,[28] Lindsay M. Morton,[3] Elizabeth A. Holly,[11] Alex Smith,[24] Lesley F. Tinker,[29] Lauren R. Teras,[30] Anne Kricker,[31] Nikolaus Becker,[32] Mark P. Purdue,[3] John J. Spinelli,[33,34] Yawei Zhang,[35] Graham G. Giles,[21,22] Paolo Vineis,[20,36] Alain Monnereau,[37,38,39] Kimberly A. Bertrand,[8,40] Demetrius Albanes,[3] Anne Zeleniuch-Jacquotte,[41,42] Attilio Gabbas,[43] Charles C. Chung,[3] Laurie Burdett,[5] Amy Hutchinson,[5] Charles Lawrence,[44] Rebecca Montalvan,[44] Liming Liang,[40,45] Jinyan Huang,[40] Baoshan Ma,[40,46] Jianjun Liu,[10] Hans-Olov Adami,[40,47] Bengt Glimelius,[48,49] Yuanqing Ye,[26] Grzegorz S. Nowakowski,[13] Ahmet Dogan,[50] Carrie A. Thompson,[51] Thomas M. Habermann,[51] Anne J. Novak,[51] Mark Liebow,[51] Thomas E. Witzig,[51] George J. Weiner,[28] Maryjean Schenk,[52] Patricia Hartge,[3] Anneclaire J. De Roos,[29,53] Wendy Cozen,[54,55] Degui Zhi,[56] Nicholas K. Akers,[2] Jacques Riby,[1,2] Martyn T. Smith,[2] Mortimer Lacher,[4] Danylo J. Villano,[4] Ann Maria,[4] Eve Roman,[24] Eleanor Kane,[24] Rebecca D. Jackson,[57] Kari E. North,[58,59] W. Ryan Diver,[30] Jenny Turner,[60,61] Bruce K. Armstrong,[31] Yolanda Benavente,[14,15] Paolo Boffetta,[62] Paul Brennan,[63] Lenka Foretova,[64] Marc Maynadie,[65] Anthony Staines,[66] James McKay,[67] Angela R. Brooks-Wilson,[68,69] Tongzhang Zheng,[35] Theodore R. Holford,[70] Saioa Chamosa,[71] Rudolph Kaaks,[32] Rachel S. Kelly,[36,40] Bodil Ohlsson,[72] Ruth C. Travis,[73] Elisabete Weiderpass,[47,74,75,76] Jacqueline Clavel,[37,38] Edward Giovannucci,[8,40,77] Peter Kraft,[40,45] Jarmo Virtamo,[78] Patrizio Mazza,[79] Pierluigi Cocco,[43] Maria Grazia Ennas,[80] Brian C.H. Chiu,[81] Joseph F. Fraumeni, Jr.,[3] Alexandra Nieters,[82,85] Kenneth Offit,[4,85] Xifeng Wu,[26,85] James R. Cerhan,[13,85] Karin E. Smedby,[83,85] Stephen J. Chanock,[3,85] and Nathaniel Rothman[3,85]

[1]Department of Epidemiology, School of Public Health and Comprehensive Cancer Center, Birmingham, AL 35233, USA; [2]Division of Environmental Health Sciences, University of California Berkeley School of Public Health, Berkeley, CA 94720, USA; [3]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, MD 20892, USA; [4]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA; [5]Cancer Genomics Research Laboratory, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Gaithersburg, MD 20877, USA; [6]Department of Medical Genetics and of Epidemiology, University Medical Center Utrecht, Utrecht 3584 CG, the Netherlands; [7]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht 3584 CX, the Netherlands; [8]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA; [9]Prince of Wales Clinical School, University of New South Wales, Sydney, NSW 2052, Australia; [10]Human Genetics, Genome Institute of Singapore, Singapore 138672, Singapore; [11]Department of Epidemiology & Biostatistics, University of California, San Francisco, San Francisco, CA 94118, USA; [12]Institute for Risk Assessment Sciences, Utrecht University, Utrecht 3508 TD, the Netherlands; [13]Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA; [14]Unit of Infections and Cancer (UNIC), Cancer Epidemiology Research Programme, Institut Catala d'Oncologia, IDIBELL, Barcelona 8907, Spain; [15]Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), Barcelona 8036, Spain; [16]Department of Cancer Etiology, City of Hope Beckman Research Institute, Duarte, CA 91030, USA; [17]Department of Hematology, Hospices Civils de Lyon, Pierre benite Cedex 69495, France; [18]Department of Hematology, Université Lyon-1, Pierre benite Cedex 69495, France; [19]Laboratoire de Biologie Moléculaire de la Cellule UMR 5239, Centre National de la Recherche Scientifique, Pierre benite Cedex 69495, France; [20]Human Genetics Foundation, Turin 10126, Italy; [21]Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, VIC 3053, Australia; [22]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Carlton, VIC 3010, Australia; [23]Department of Epidemiology Research, Division of Health Surveillance and Research, Statens Serum Institut, Copenhagen 2300, Denmark; [24]Department of Health Sciences, University of York, York YO10 5DD, UK; [25]Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA; [26]Department of Epidemiology, M.D. Anderson Cancer Center, Houston, TX 77030, USA; [27]Department of Hematology, Centre Léon Bérard, Lyon 69008, France; [28]Department of Internal Medicine, Carver College of Medicine, The University of Iowa, Iowa City, IA 52242, USA; [29]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98117, USA; [30]Epidemiology Research Program, American Cancer Society, Atlanta, GA 30303, USA; [31]Sydney School of Public Health, The University of Sydney, Sydney, NSW 2006, Australia; [32]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Baden-Württemberg 69120, Germany; [33]Cancer Control Research, BC Cancer Agency, Vancouver, BC V5Z 1L3, Canada; [34]School of Population and Public Health, University of British Columbia, Vancouver, BC V6T 1Z3, Canada; [35]Department of Environmental Health Sciences, Yale School of Public Health, New Haven, CT 06520, USA; [36]MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London W2 1PG, UK; [37]Environmental Epidemiology of Cancer Group, Inserm, Centre for Research in Epidemiology and Population Health (CESP), U1018, Villejuif Cedex 94807, France; [38]UMRS 1018, Université Paris Sud, Villejuif Cedex 94807, France; [39]Registre des hémopathies malignes de la Gironde, Institut Bergonié, Bordeaux Cedex 33076, France; [40]Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA; [41]Department of Population Health, New York University School of Medicine, New York, NY 10016, USA; [42]Cancer Institute, New York University School of Medicine, New York, NY 10016, USA; [43]Department of Public Health, Clinical and Molecular Medicine, University of Cagliari, Monserrato, Cagliari 09042, Italy; [44]Health Studies Sector, Westat, Rockville, MD 20850, USA; [45]Department of Biostatistics, Harvard School of Public

Genome-wide association studies (GWASs) of follicular lymphoma (FL) have previously identified human leukocyte antigen (HLA) gene variants. To identify additional FL susceptibility loci, we conducted a large-scale two-stage GWAS in 4,523 case subjects and 13,344 control subjects of European ancestry. Five non-HLA loci were associated with FL risk: 11q23.3 (rs4938573, $p = 5.79 \times 10^{-20}$) near *CXCR5*; 11q24.3 (rs4937362, $p = 6.76 \times 10^{-11}$) near *ETS1*; 3q28 (rs6444305, $p = 1.10 \times 10^{-10}$) in *LPP*; 18q21.33 (rs17749561, $p = 8.28 \times 10^{-10}$) near *BCL2*; and 8q24.21 (rs13254990, $p = 1.06 \times 10^{-8}$) near *PVT1*. In an analysis of the HLA region, we identified four linked HLA-DRβ1 multiallelic amino acids at positions 11, 13, 28, and 30 that were associated with FL risk ($p_{omnibus} = 4.20 \times 10^{-67}$ to 2.67×10^{-70}). Additional independent signals included rs17203612 in HLA class II (odds ratio [$OR_{per\text{-}allele}$] = 1.44; $p = 4.59 \times 10^{-16}$) and rs3130437 in HLA class I ($OR_{per\text{-}allele} = 1.23$; $p = 8.23 \times 10^{-9}$). Our findings further expand the number of loci associated with FL and provide evidence that multiple common variants outside the HLA region make a significant contribution to FL risk.

Follicular lymphoma (FL [MIM 613024]) is a common B cell malignancy characterized by a variable indolent clinical course that can take decades to manifest and, in some cases, can be followed by transformation to aggressive diffuse large B cell lymphoma (DLBCL).[1,2] The previous genome-wide association studies (GWASs) of relatively small sample sizes have revealed FL susceptibility loci in the human leukocyte antigen (HLA) class I and class II regions on 6p21.32-33.[3–7] To identify new FL susceptibility loci, we genotyped 2,301 FL case subjects and 2,854 control subjects of European descent from 22 studies (NCI FL GWAS) as part of a larger initiative using the Illumina OmniExpress Beadchip (Table S1; Figure S1 available online). All studies obtained informed consent from participants and approval from the respective Institutional Review Boards for this study. Cases were ascertained from cancer registries, clinics, or hospitals or through self-report verified by medical and pathology reports (Table S1). The phenotype information for all cases was reviewed centrally at the International Lymphoma Epidemiology Consortium (InterLymph) Data Coordinating Center, and cases were classified according to the proposed scheme by the InterLymph Pathology Working Group based on the World Health Organization (WHO) classification (2008) (Table S1). Genotypes were called using Illumina GenomeStudio software, and quality-control duplicates showed >99% concordance. All initial data analyses and management were conducted using the Genotyping Library and Utilities (GLU), and extensive quality-control metrics were applied to the data. Specifically, monomorphic SNPs and SNPs with call rates <93% were removed, and samples with call rates ≤93%, mean heterozygosity <0.25 or >0.33 based on the autosomal SNPs, or gender discordance (>5% heterozygosity on the X chromosome for males and <20% heterozygosity on the X chromosome for females) were excluded. Unexpected duplicates (>99.9% concordance) and first-degree relatives on the basis of identity-by-descent sharing with Pi-hat >0.40 were removed. Ancestry was assessed using the GLU struct.admix module, and participants with <80% European ancestry were also excluded (Figure S2). After these quality-control steps, 94% of the participants and 611,844 SNPs remained for analysis (Tables S2 and S3). Genotype data previously generated on the Illumina Omni2.5 BeadChip[8] from an additional 3,536 control subjects from 3 of the 22 studies (ATBC, CPSII, and PLCO) were

Health, Boston, MA 02115, USA; [46]College of Information Science and Technology, Dalian Maritime University, Dalian, Liaoning Province 116026, China; [47]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm 17177, Sweden; [48]Department of Oncology and Pathology, Karolinska Institutet, Karolinska University Hospital Solna, Stockholm 17176, Sweden; [49]Department of Radiology, Oncology and Radiation Science, Uppsala University, Uppsala 75105, Sweden; [50]Departments of Laboratory Medicine and Pathology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA; [51]Department of Medicine, Mayo Clinic, Rochester, MN 55905, USA; [52]Department of Family Medicine and Public Health Sciences, Wayne State University, Detroit, MI 48201, USA; [53]Department of Environmental and Occupational Health, Drexel University School of Public Health, Philadelphia, PA 19104, USA; [54]Department of Preventive Medicine, USC Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA; [55]Norris Comprehensive Cancer Center, USC Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA; [56]Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL 35233, USA; [57]Division of Endocrinology, Diabetes and Metabolism, The Ohio State University, Columbus, OH 43210, USA; [58]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA; [59]Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA; [60]Department of Anatomical Pathology, Australian School of Advanced Medicine, Macquarie University, Sydney, NSW 2109, Australia; [61]Department of Histopathology, Douglass Hanly Moir Pathology, Macquarie Park, NSW 2113, Australia; [62]The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [63]Group of Genetic Epidemiology, Section of Genetics, International Agency for Research on Cancer, Lyon 69372, France; [64]Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute and MF MU, Brno 656 53, Czech Republic; [65]EA 4184, Registre des Hémopathies Malignes de Côte d'Or, University of Burgundy and Dijon University Hospital, Dijon 21070, France; [66]School of Nursing and Human Sciences, Dublin City University, Dublin 9, Ireland; [67]Genetic Cancer Susceptibility Group, Section of Genetics, International Agency for Research on Cancer, Lyon 69372, France; [68]Genome Sciences Centre, BC Cancer Agency, Vancouver, BC V5Z 1L3, Canada; [69]Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, BC V5A 1S6, Canada; [70]Department of Biostatistics, Yale School of Public Health, New Haven, CT 06520, USA; [71]Health Department, BioDonostia Research Institute, Basque Region 20014, Spain; [72]Department of Clinical Sciences, Division of Internal Medicine, Skåne University Hospital, Lund University, Malmö 205 02, Sweden; [73]Cancer Epidemiology Unit, University of Oxford, Oxford OX3 7LF, UK; [74]Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, Breivika 9037, Norway; [75]Cancer Registry of Norway, Oslo 0304, Norway; [76]Department of Genetic Epidemiology, Folkhalsan Research Center, Helsinki 00250, Finland; [77]Department of Nutrition, Harvard School of Public Health, Boston, MA 02115, USA; [78]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki 00271, Finland; [79]Department of Hematology, Ospedale Nord, Taranto 74100, Italy; [80]Department of Biomedical Science, University of Cagliari, Monserrato, Cagliari 09042, Italy; [81]Department of Health Studies, University of Chicago, Chicago, IL 60637, USA; [82]Center for Chronic Immunodeficiency, University Medical Center Freiburg, Freiburg, Baden-Württemberg 79108, Germany; [83]Department of Medicine Solna, Karolinska Institutet, Stockholm 17176, Sweden
[84]These authors contributed equally to this work
[85]These authors contributed equally to this work in a supervisory role
*Correspondence: cskibola@uab.edu
http://dx.doi.org/10.1016/j.ajhg.2014.09.004.

Table 1. Association Results for Loci and SNPs Associated with Risk of Follicular Lymphoma

Chr	Nearest Gene(s)	SNP	Position[a]	Risk Allele[b]	Other Allele	RAF[c]	Stage	OR	(95% CI)	p	Phet[d]	I^2[e]
Known Locus												
6p21.32	HLA region	rs12195582 (rs115374828)	32444544	T	C	0.465	NCI	1.88	(1.74–2.02)	3.26×10^{-58}	–	–
						0.498	previous GWAS	1.55	(1.33–1.80)	1.10×10^{-8}		
						0.435	replication	1.75	(1.60–1.90)	1.17×10^{-37}		
						–	combined	1.78	(1.69–1.88)	5.36×10^{-100}	2.75×10^{-1}	19.56
Genome-wide Significant Loci												
11q23.3	*CXCR5*	rs4938573	118741842	C	T	0.204	NCI	1.30	(1.19–1.43)	5.97×10^{-9}	–	–
						0.193	previous GWAS	1.37	(1.14–1.64)	0.0008		
						0.188	replication	1.39	(1.25–1.54)	3.17×10^{-10}		
						–	combined	1.34	(1.26–1.43)	5.79×10^{-20}	7.69×10^{-1}	0.00
11q24.3	*ETS1*	rs4937362	128492739	T	C	0.456	NCI	1.16	(1.08–1.25)	7.01×10^{-5}	–	–
						0.465	previous GWAS	1.33	(1.16–1.54)	5.90×10^{-5}		
						0.467	replication	1.17	(1.08–1.28)	0.0002		
						–	combined	1.19	(1.13–1.25)	6.76×10^{-11}	7.52×10^{-1}	0.00
3q28	*LPP*	rs6444305	188299902	G	A	0.276	NCI	1.16	(1.08–1.27)	0.0002	–	–
						0.269	previous GWAS	1.30	(1.06–1.59)	0.01		
						0.281	replication	1.25	(1.14–1.37)	2.21×10^{-6}		
						–	combined	1.21	(1.14–1.28)	1.10×10^{-10}	4.42×10^{-1}	0.00
18q21.33	*BCL2*	rs17749561	60783211	G	A	0.910	NCI	1.43	(1.25–1.61)	2.18×10^{-7}	–	–
						0.908	previous GWAS	1.23	(0.96–1.57)	1.10×10^{-1}		
						0.905	replication	1.28	(1.10–1.49)	0.002		
						–	combined	1.34	(1.22–1.47)	8.28×10^{-10}	5.43×10^{-2}	49.37
8q24.21	*PVT1*	rs13254990	129076451	T	C	0.315	NCI	1.20	(1.11–1.30)	8.39×10^{-6}	–	–
						0.307	previous GWAS	1.15	(0.98–1.34)	0.08		
						0.315	replication	1.16	(1.06–1.27)	0.001		
						–	combined	1.18	(1.11–1.24)	1.06×10^{-8}	6.99×10^{-1}	0.00
Suggestive Loci												
17q25.3	*C17orf62*	rs3751913	80405552	C	T	0.121	NCI	1.25	(1.11–1.39)	0.0001	–	–
						0.126	previous GWAS	1.42	(1.16–1.75)	0.0008		
						0.121	replication	1.14	(1.01–1.29)	0.04		
						–	combined	1.23	(1.14–1.33)	2.24×10^{-7}	2.59×10^{-1}	21.50
3q13.33	*CD86*	rs2681416	121817613	A	G	0.311	NCI	1.24	(1.15–1.35)	6.73×10^{-8}	–	–
						0.305	previous GWAS	1.15	(0.99–1.34)	0.06		
						0.329	replication	1.06	(0.97–1.15)	0.23		
						–	combined	1.16	(1.09–1.22)	2.33×10^{-7}	5.54×10^{-4}	72.83
18q12.3	*SLC14A2*	rs11082438	42865210	G	T	0.936	NCI	1.39	(1.18–1.61)	4.65×10^{-5}	–	–
						0.941	previous GWAS	1.46	(1.07–1.99)	0.02		
						0.935	replication	1.22	(1.02–1.46)	0.03		
						–	combined	1.33	(1.19–1.48)	4.01×10^{-7}	9.26×10^{-1}	0.00

[a]Position according to human reference NCBI37/hg19.
[b]Allele associated with an increased risk of FL.
[c]Risk allele frequency in controls.
[d]Cochran's Q test heterogeneity p value.
[e]I^2 heterogeneity index.

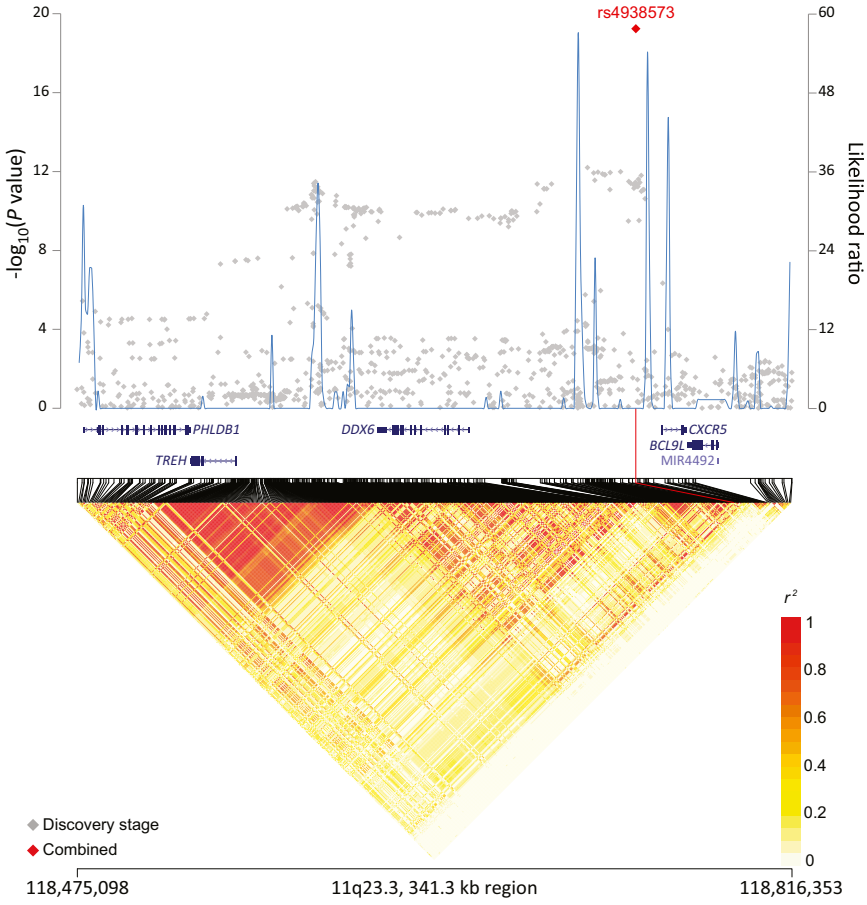


Figure 1. Regional Plots of the FL-Associated Locus rs4938573 in 11q23.3
Figure shows the association results from the NCI FL GWAS and stages 1 and 2 combined (red diamond), recombination hotspots, and LD plots.

also included, resulting in a total of 2,142 FL case subjects and 6,221 control subjects for analysis (NCI FL GWAS; Table S4).

To evaluate population substructure, a principal components analysis was conducted using the GLU struct.pca module. Plots of the top principal components are shown in Figure S3. Association testing was conducted assuming a log-additive genetic model adjusted for age, sex, and significant principal components. A quantile-quantile plot of the association results revealed an enrichment of SNPs with small p values even after removal of all SNPs in the HLA region, which has been previously reported to be associated with FL (lambda = 1.018, Figure S4). In addition to the HLA region, one locus on 11q23.3 reached genome-wide statistical significance ($p < 5 \times 10^{-8}$) (Figure S5).

To increase power to detect associations in stage 1, we added data on 586 FL case subjects and 1,537 control subjects from two independent previously published GWASs (UCSF2[4] and SCALE[3]) to the newly genotyped NCI FL GWAS (Tables S1 and S4; Figure S1). Because different genotyping platforms were used (Table S2), we imputed all three GWASs (NCI, UCSF2, SCALE) using the 1000 Genomes Project (1kGP) v.3 (March 2012 release) reference panel[9] and IMPUTE2.[10] The genotype data underwent rigorous quality control filters before imputation (Table S2), and association testing was conducted separately for each study using SNPTEST v.2 adjusted for age, sex, and significant principal components.

Association results from the NCI FL GWAS and the two previously genotyped GWASs (totaling 2,728 case subjects, 7,758 control subjects in stage 1) were analyzed in a meta-analysis using a fixed-effects inverse-variance method based on the β estimates and standard errors from each study. Only SNPs with information scores >0.3 were included in the meta-analysis. In the stage 1 meta-analysis, we identified three non-HLA loci (11q23.3, 11q24.3, and 3q13.33) that reached genome-wide significance ($p < 5 \times 10^{-8}$). To confirm these loci and discover additional loci, 11 non-HLA SNPs with $p < 5 \times 10^{-6}$ from the stage 1 meta-analysis were chosen for replication in stage 2. Only SNPs with a MAF > 1% were considered for replication, and no SNPs were taken forward for replication in regions where they appeared to be singletons or obvious artifacts. Stage 2 replication was undertaken in a new set of 1,795 FL case subjects and 5,586 control subjects, which included 119 case subjects and 349 control subjects from another GWAS (UCSF1/NHS) genotyped on the OmniExpress microarray and imputed using IMPUTE2[10] and the 1kGP data,[9] and 1,676 cases and 5,237 controls with de novo genotyping (Tables S1, S2, and S4). All 11 SNPs were either directly genotyped or had a high imputation information score (average information score = 0.92). Genotyping of these 11 SNPs by TaqMan (Applied Biosystems) in 470 subjects from the NCI GWAS yielded >88.9% concordance with the imputed dosages (median concordance = 99.6%), indicating that imputation accuracy was high. Association testing was conducted for each study using either GLU (de novo genotyping) or SNPTEST (UCSF1/NHS), adjusting for relevant factors.

Results from the stage 1 and 2 studies were then meta-analyzed using a fixed effects model. In the combined meta-analysis, we found five non-HLA loci that achieved genome-wide significance ($p < 5 \times 10^{-8}$) at 11q23.3 (rs4938573, $p = 5.79 \times 10^{-20}$), 11q24.3 (rs4937362, $p = 6.76 \times 10^{-11}$), 3q28 (rs6444305, $p = 1.10 \times 10^{-10}$),

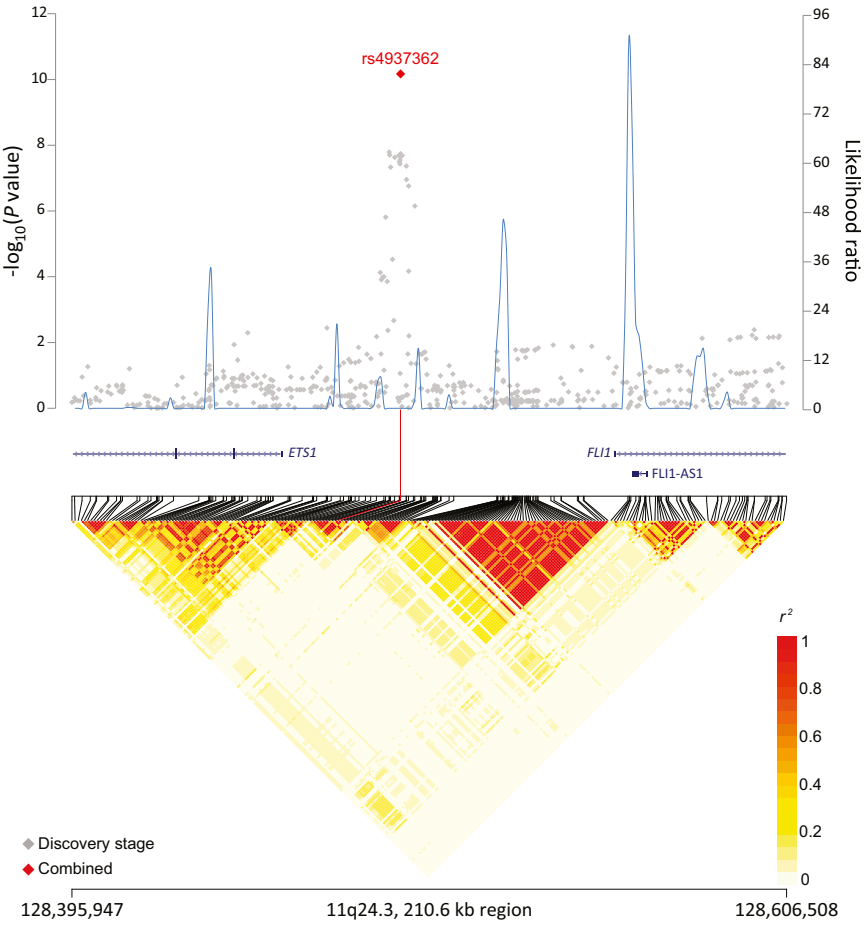


Figure 2. Regional Plots of the FL-Associated Locus rs4937362 in 11q24.3
Figure shows the association results from the NCI NHL GWAS and stages 1 and 2 combined (red diamond), recombination hotspots, and LD plots.

18q21.33 (rs17749561, $p = 8.28 \times 10^{-10}$), and 8q24.21 (rs13254990, $p = 1.06 \times 10^{-8}$); and three suggestive loci ($p < 5 \times 10^{-7}$) at 17q25.3 (rs3751913, $p = 2.24 \times 10^{-7}$), 3q13.33 (rs2681416, $p = 2.33 \times 10^{-7}$), and 18q12.3 (rs11082438, $p = 4.01 \times 10^{-7}$) (Table 1). Two of the five loci that reached genome-wide significance in the stage 1 and 2 meta-analysis (11q23.3 and 11q24.3) were genome-wide significant in the stage 1 meta-analysis and were robustly replicated in stage 2 ($p = 3.17 \times 10^{-10}$ and $p = 0.0002$, respectively). The remaining three loci achieved genome-wide significance after inclusion of the stage 2 data and therefore would benefit from further validation in other independent samples.

rs4938573 at 11q23.3 maps 12.6 kb upstream of the chemokine (c-x-c motif) receptor 5 gene (*CXCR5* [MIM 601613]) (Figure 1). The 11q24.3 locus marked by rs4937362 ($p = 6.76 \times 10^{-11}$) is approximately 35 kb upstream of v-ets avian erythroblastosis virus E26 oncogene homolog 1 (*ETS1* [MIM 164720]) (Figure 2). The 3q28 locus marked by rs6444305 maps to a region that overlaps the LIM domain containing preferred translocation partner in lipoma (*LPP* [MIM 600700]) and is 836.4 kb upstream of *BCL6* (MIM 109565) (Figure 3). rs17749561 in 18q21.33 is located 7.4 kb downstream of the antiapoptotic oncogene, B cell CLL/lymphoma 2 (*BCL2* [MIM 151430]) (Figure 4); and rs13254990 at 8q24.21 maps near the oncogene, plasmacytoma variant translocation 1 gene (*PVT1* [MIM 165140]) (Figure 5). Characteristics of these loci are presented in Table S5. The suggestive SNP rs3751913 is in chromosome 17 opening reading frame 62 (*C17orf62*); rs2681416 is in CD86 molecule (*CD86*) (MIM 601020); and rs11082438 is in solute carrier 14A2 (*SLC14A2* [MIM 601611]) (Table 1, Figure S6). Using the Cochran's Q test and by estimating the I^2 heterogeneity index, no substantial heterogeneity was observed among the studies for any SNP ($p_{heterogeneity} \geq 0.05$) except for the suggestive locus, rs2681416 at 3q13.33 (Table 1). Although the p value for heterogeneity for rs13254990 was borderline significant, all of the effect estimates for the individual studies were above 1.0.

To explore potential functional roles for associated SNPs and their surrogates ($r^2 > .80$) and to assess the B cell-specific chromatin dynamics of regions overlapping with the associated SNPs, we conducted HaploReg[11] and ChroMoS analyses.[12,13] Here we found that three loci, 11q23.3, 3q13.33, and 8q24.21, were annotated as overlapping enhancers in the lymphoblastoid cell line GM12878,[14] suggesting that our GWAS signals map to variants that overlap within regions of active chromatin state in B cells (Table S6; Figure S7). However, an expression quantitative trait loci (eQTL) analysis using publicly available RNA sequencing data on lymphoblastoid cell lines (available from the Gene Expression Omnibus [GEO] repository under accession number GSE16921) yielded no notable (FDR < 0.05) associations of the selected SNPs with gene expression levels. Additional analysis using microarray data (GEO accession number GSE8052) did not reveal any significant eQTL associations for the genome-wide significant loci, although the suggestive SNP, rs3751913, was associated with *C17orf62* expression (data not shown). Thus, further work is needed to identify and characterize the biological basis of these FL susceptibility alleles.

Consistent with previous smaller reports, the strongest effects on FL risk were observed in the HLA region at 6p21.32-33, where 8,104 SNPs achieved genome-wide

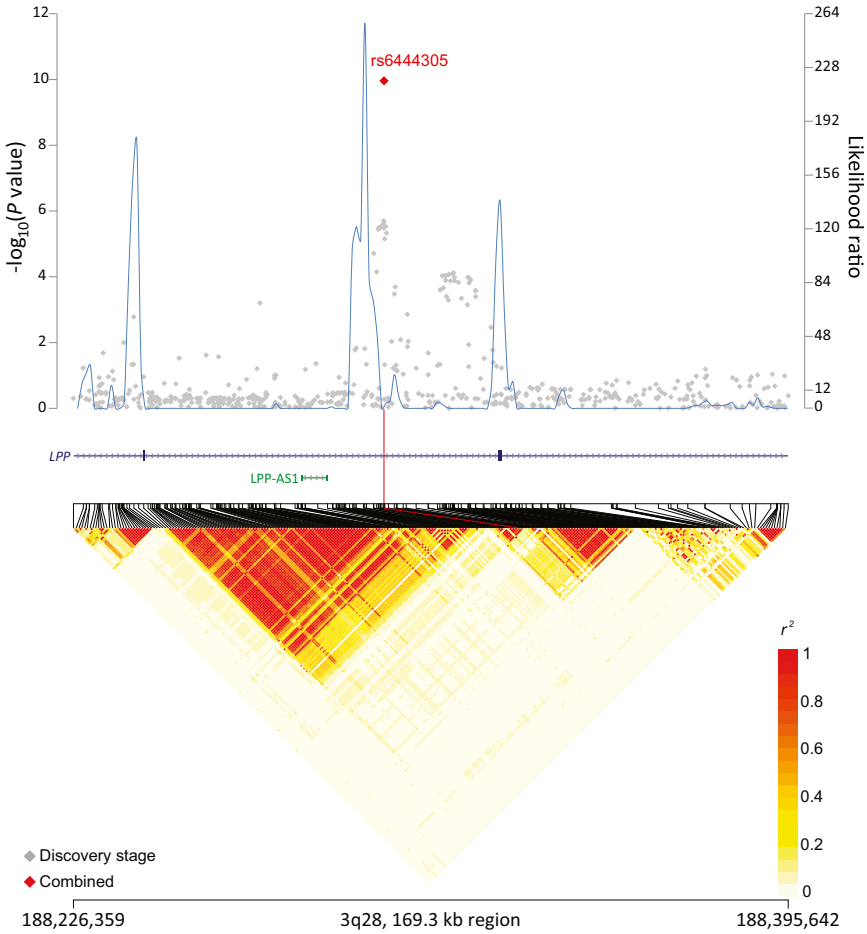


Figure 3. Regional Plots of the Associated Locus rs6444305 in 3q28
Figure shows the association results from the NCI FL GWAS and stages 1 and 2 combined (red diamond), recombination hotspots, and LD plots.

significance ($p < 5 \times 10^{-8}$) in the stage 1 meta-analysis (Figure S8). One top SNP, rs12195582, was carried forward for replication in stage 2 and reached a combined $p = 5.36 \times 10^{-100}$ in stages 1+2 (Table 1). To further refine the association of HLA variants with FL risk and determine whether specific coding variants within HLA genes contributed to the diverse association signals, we imputed classical HLA alleles and amino acids (AAs) at seven loci (*HLA-A* [MIM 142800], *HLA-B* [MIM 142830], *HLA-C* [MIM 142840], *HLA-DQA1* [MIM 146880], *HLA-DQB1* [MIM 604305], *HLA-DRB1* [MIM 142857], and *HLA-DPB1* [MIM 142858]) on the four GWAS data sets from stages 1+2 (NCI, USCF2, SCALE, UCSF1/NHS) using SNP2HLA[15] and a reference panel from the Type 1 Diabetes Genetics Consortium (T1DGC) consisting of genotype data from 5,225 individuals of European descent that were typed for classical HLA alleles. The imputation accuracy of HLA types was high (>95.23%) when compared to HLA sequencing data on a subset of NCI and UCSF2 samples scanned as part of this study.[16,17] Due to the limited number of SNPs (7,253) in the T1DGC reference set, imputation of HLA SNPs was conducted with IMPUTE2 and the 1kGP reference set. A total of 68,488 SNPs, 201 classical HLA alleles (two- and four-digit resolution), and 1,038 AA markers including 103 AA positions that were "multiallelic" with three to six different residues present at each position, were successfully imputed (information score > 0.3 for SNPs or $r^2 > 0.3$ for alleles and AAs) and available for downstream analysis. Association testing was conducted using PLINK,[18] where multiallelic markers were analyzed as binary markers (e.g., allele present or absent). A meta-analysis was conducted where we tested SNPs, HLA alleles, and AAs across the HLA region for association to FL. Among the imputed AAs and HLA alleles tested, the top associated signal mapped to a DRβ1 AA at position 28 that carries three possible amino acids: Glu, Asp, and His. Asp was associated with low ($OR = 0.53$; $p = 6.1 \times 10^{-72}$) and Glu with high ($OR = 1.86$; $p = 7.99 \times 10^{-69}$) FL risk (Table S7). Global omnibus tests of position 28 ($2.49 \times 10^{-67} \leq p \leq 3.84 \times 10^{-67}$) and other nearby DRβ1 AA positions at 11, 13, and 30 yielded statistically similar associations with FL risk (Table S9). These results support the previously reported association between FL and DRβ1 position 13 in a small study of Europeans.[19] However, due to the high LD between positions 11, 13, 28, and 30, we were unable to determine the significance of one position at the exclusion of the other through reciprocal conditional analyses. The most significant imputed two- or four-digit HLA allele in our analysis was *DRB1*01* ($OR = 1.85$; $p = 2.57 \times 10^{-42}$) (Table S7), encoded by Glu28, Cys30, Phe13, and Leu11 (Table S9). An association with FL risk was found for *HLA-DRB1*07:01* that is also encoded by residues at 11, 13, 28, and 30 ($p = 1.59 \times 10^{-20}$) (Table S9). Positions 11, 13, 28, and 30 reside in the middle of the HLA-DR heterodimer molecule in the peptide binding cleft (Figure S9) that specifically impact binding pockets 4, 6, and 7. These are key peptide binding anchors in DRβ1[20] that influence binding preferences of alleles,[21] suggesting an important role for DRβ1 peptide presentation in follicular lymphomagenesis.

To identify independent HLA variants controlling for DRβ1 28 (used as a surrogate for the 11, 13, 28, and 30 group), we included all genotyped and imputed HLA SNPs, AAs, and alleles in a forward stepwise analysis. The most significant variant after controlling for DRβ1 28 was

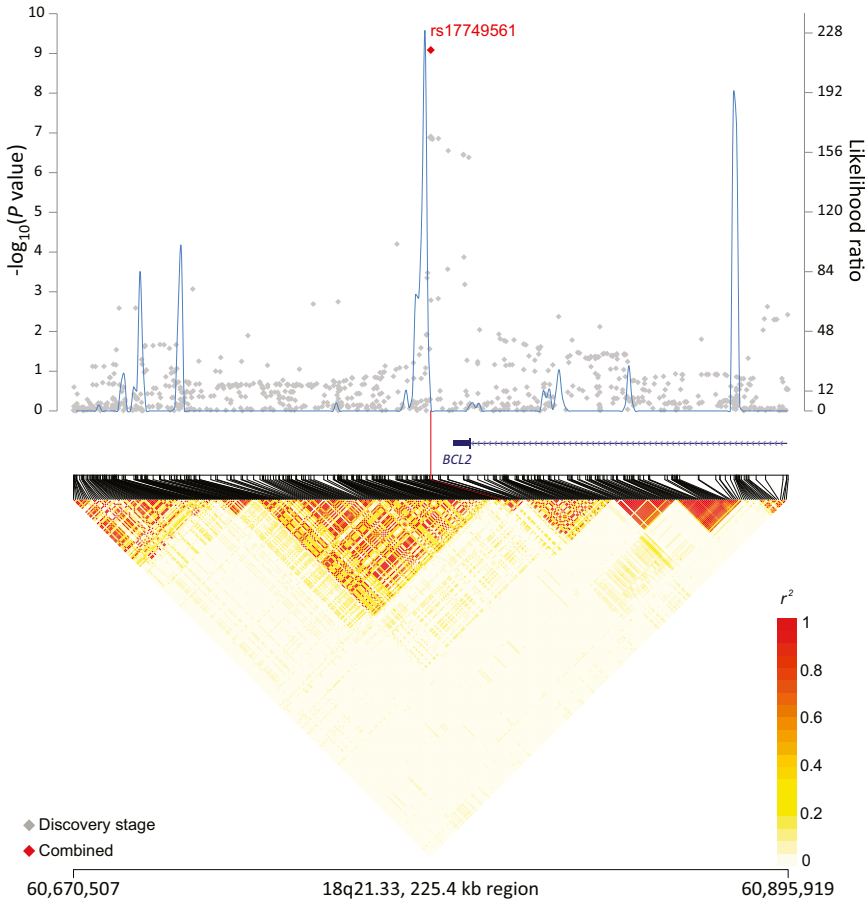


Figure 4. Regional Plots of the Associated Locus rs17749561 in 18q21.33
Figure shows the association results from the NCI FL GWAS and stages 1 and 2 combined (red diamond), recombination hotspots, and LD plots.

rs17203612 ($p = 4.59 \times 10^{-16}$), an intergenic SNP 39.2 kb and 99.7 kb downstream of *HLA-DRA* (MIM 142860) and *HLA-DRB1*, respectively (Figure 6; Table S10). A conditional analysis on DRβ1 28 and rs17203612 revealed that the next most statistically significant variant was rs3130437 ($p = 8.23 \times 10^{-9}$) located 15.6 kb downstream of *HLA-C* in HLA class I (Figure 6; Table S10). After controlling for DRβ1 28, rs17203612, and rs3130437, no additional signals with $p < 5 \times 10^{-8}$ were observed (Figure 6). Of note, we did see a residual signal ($p = 8.18 \times 10^{-6}$) at the functionally relevant DPβ1 Glu84 position,[22] a reported risk locus for Hodgkin lymphoma.[23] A conditional analysis of DRβ1 28, rs17203612, and rs3130437 eliminated the majority of residual effects for the previously reported HLA SNPs and alleles associated with FL (Table S11).

We conducted a series of preliminary bioinformatics analyses to explore the potential functional relevance of rs17203612 and rs3130437 using publicly available RNA sequencing expression and methylation data and found significant (FDR < 0.05) gene expression and methylation differences associated with rs17203612- and rs3130437-linked SNPs (Tables S12 and S13). Specifically, we found significant gene expression changes associated with rs12194148, a proxy for rs17203612, in class II (*HLA-DRB5* [MIM 604776], *HLA-DRB6*, *HLA-DRB1*, *HLA-DQB1*, *HLA-DQB2* [MIM 615161], *HLA-DQA1*, *HLA-DQA2* [MIM 613503], *BTNL2* [MIM 606000], *C6orf25*); and with rs3130439, a proxy for rs3130437, in HLA class I (*PSORS1C2*, *PSORS1C3*, *DPCR1* [MIM 613928]) (Table S12). Of note, ten of the rs17203612-linked SNPs that showed correlation with higher *HLA-DQB1* expression also showed correlation with lower *HLA-DQB1* methylation levels (Table S12) that further supports the potential role of HLA class II FL-associated SNPs in *HLA-DQB1* regulation.[24,25] Additional eQTL analyses using microarray data also suggested potential eQTL associations with *HLA-C*, *TCF19* (MIM 600912), and *HLA-B* expression (Table S14). However, we did not observe significant enrichment of particular regulatory markers within these associated regions, although overlap with some regulatory signals was observed (Table S15).

In summary, our study identified five non-HLA susceptibility alleles that were robustly associated with FL risk. Moreover, our work highlights the important role of HLA structural variants and regulatory SNPs in the etiology of FL, advances the catalog of HLA and non-HLA genetic variants associated with FL risk, and provides further evidence for a role of DRβ1 peptide presentation in FL. Functional studies will be required to elucidate the biological basis of these loci and to determine their role in follicular lymphomagenesis.

Supplemental Data

Supplemental Data include 9 figures, 15 tables, and Supplemental Acknowledgments and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2014.09.004.

Acknowledgments

The overall FL GWAS project was supported by the intramural program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH. A full list of Supplemental Acknowledgements is provided online.

Received: July 17, 2014
Accepted: September 10, 2014
Published: October 2, 2014

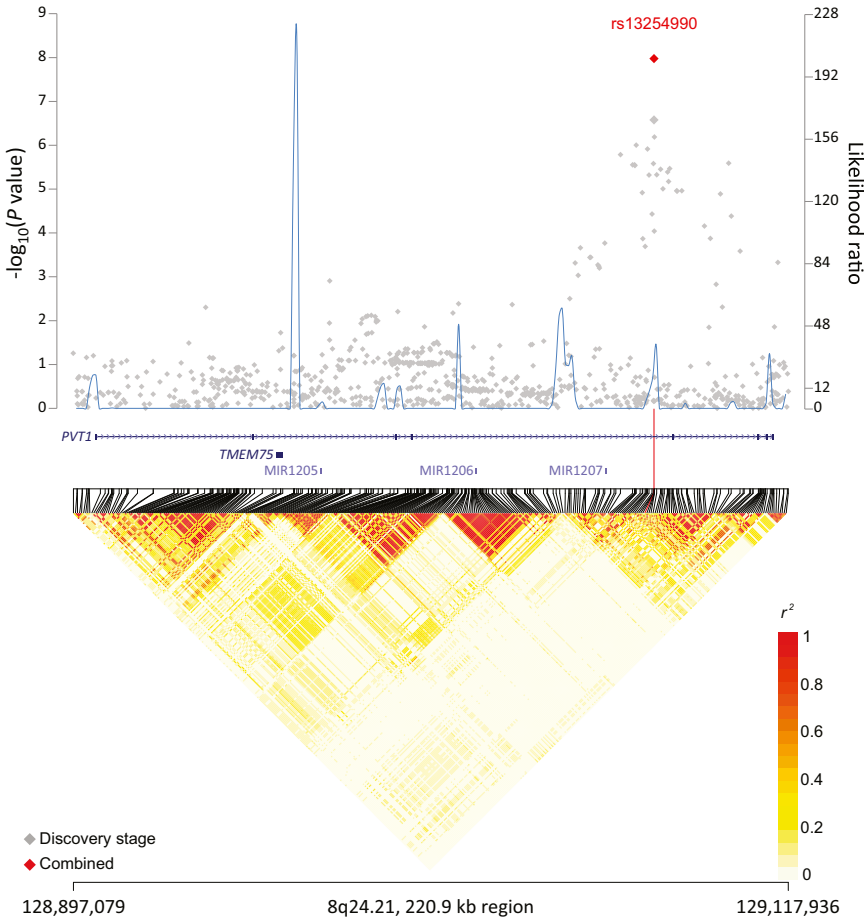


Figure 5. Regional Plots of the Associated Locus rs13254990 in 8q24.21
Figure shows the association results from the NCI FL GWAS and stages 1 and 2 combined (red diamond), recombination hotspots, and LD plots.

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, http://browser.1000genomes.org
ChroMoS, http://epicenter.immunbio.mpg.de/services/chromos
Gene Expression Omnibus (GEO), http://www.ncbi.nlm.nih.gov/geo/
glu-genetics, https://code.google.com/p/glu-genetics/
HaploReg, http://www.broadinstitute.org/mammals/haploreg/haploreg.php
IMPUTE2, http://mathgen.stats.ox.ac.uk/impute/impute_v2.html
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/
PLINK, http://pngu.mgh.harvard.edu/~purcell/plink/
SNP2HLA, https://www.broadinstitute.org/mpg/snp2hla/
snptest, https://mathgen.stats.ox.ac.uk/genetics_software/snptest/old/snptest.html

References

1. Link, B.K., Maurer, M.J., Nowakowski, G.S., Ansell, S.M., Macon, W.R., Syrbu, S.I., Slager, S.L., Thompson, C.A., Inwards, D.J., Johnston, P.B., et al. (2013). Rates and outcomes of follicular lymphoma transformation in the immunochemotherapy era: a report from the University of Iowa/MayoClinic Specialized Program of Research Excellence Molecular Epidemiology Resource. J. Clin. Oncol. *31*, 3272–3278.
2. Tan, D., Horning, S.J., Hoppe, R.T., Levy, R., Rosenberg, S.A., Sigal, B.M., Warnke, R.A., Natkunam, Y., Han, S.S., Yuen, A., et al. (2013). Improvements in observed and relative survival in follicular grade 1-2 lymphoma during 4 decades: the Stanford University experience. Blood *122*, 981–987.
3. Smedby, K.E., Foo, J.N., Skibola, C.F., Darabi, H., Conde, L., Hjalgrim, H., Kumar, V., Chang, E.T., Rothman, N., Cerhan, J.R., et al. (2011). GWAS of follicular lymphoma reveals allelic heterogeneity at 6p21.32 and suggests shared genetic susceptibility with diffuse large B-cell lymphoma. PLoS Genet. *7*, e1001378.
4. Conde, L., Halperin, E., Akers, N.K., Brown, K.M., Smedby, K.E., Rothman, N., Nieters, A., Slager, S.L., Brooks-Wilson, A., Agana, L., et al. (2010). Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. Nat. Genet. *42*, 661–664.
5. Skibola, C.F., Bracci, P.M., Halperin, E., Conde, L., Craig, D.W., Agana, L., Iyadurai, K., Becker, N., Brooks-Wilson, A., Curry, J.D., et al. (2009). Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. Nat. Genet. *41*, 873–875.
6. Vijai, J., Kirchhoff, T., Schrader, K.A., Brown, J., Dutra-Clarke, A.V., Manschreck, C., Hansen, N., Rau-Murthy, R., Sarrel, K., Przybylo, J., et al. (2013). Susceptibility loci associated with specific and shared subtypes of lymphoid malignancies. PLoS Genet. *9*, e1003220.
7. Skibola, C.F., Conde, L., Foo, J.N., Riby, J., Humphreys, K., Sillé, F.C., Darabi, H., Sanchez, S., Hjalgrim, H., Liu, J., et al.

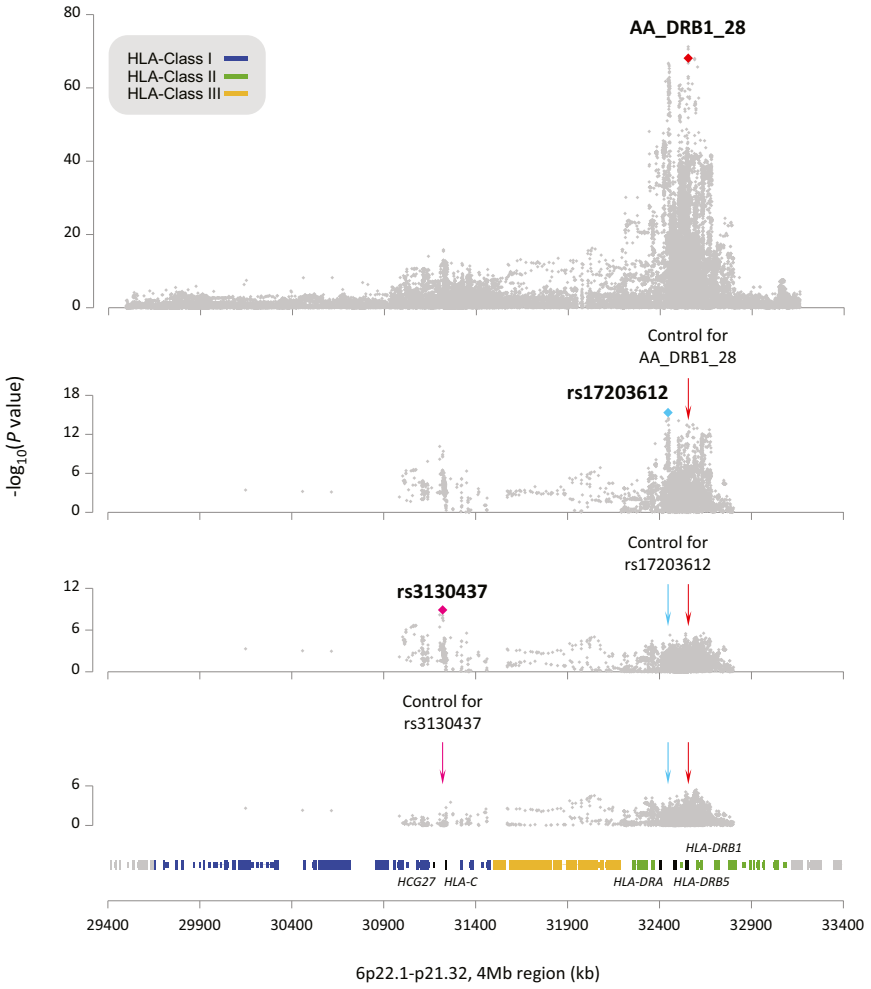


Figure 6. Sequential Conditioned Association Analysis in the HLA Region at 6p22.1-21.32: 29,400–33,400 kb
Each gray diamond represents the p value from the meta-analysis of the four GWASs. Among all the AAs and HLA alleles tested, the top associated signal mapped to the AA DRβ1 at position 28 (top). After conditioning on DRβ1 28, rs17203612 in the HLA class II region was the marker with the highest association (second from top). Further analysis conditioning on both signals revealed rs3130437 in HLA class I as the most significant associated marker (second from bottom). No additional genome-wide significant signals were observed after controlling for the effects of DRβ1 28, rs17203612, and rs3130437 (bottom). Plots derived using genome assembly hg19.

(2012). A meta-analysis of genome-wide association studies of follicular lymphoma. BMC Genomics *13*, 516.

8. Wang, Z., Jacobs, K.B., Yeager, M., Hutchinson, A., Sampson, J., Chatterjee, N., Albanes, D., Berndt, S.I., Chung, C.C., Diver, W.R., et al. (2012). Improved imputation of common and uncommon SNPs with a new reference set. Nat. Genet. *44*, 6–7.
9. Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., and McVean, G.A.; 1000 Genomes Project Consortium (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061–1073.
10. Howie, B.N., Donnelly, P., and Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. *5*, e1000529.
11. Ward, L.D., and Kellis, M. (2012). HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. *40* (Database issue), D930–D934.
12. Barenboim, M., and Manke, T. (2013). ChroMoS: an integrated web tool for SNP classification, prioritization and functional interpretation. Bioinformatics *29*, 2197–2198.
13. Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43–49.
14. Consortium, E.P., Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M.; ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.
15. Jia, X., Han, B., Onengut-Gumuscu, S., Chen, W.M., Concannon, P.J., Rich, S.S., Raychaudhuri, S., and de Bakker, P.I. (2013). Imputing amino acid polymorphisms in human leukocyte antigens. PLoS ONE *8*, e64683.
16. Wang, S.S., Abdou, A.M., Morton, L.M., Thomas, R., Cerhan, J.R., Gao, X., Cozen, W., Rothman, N., Davis, S., Severson, R.K., et al. (2010). Human leukocyte antigen class I and II alleles in non-Hodgkin lymphoma etiology. Blood *115*, 4820–4823.
17. Skibola, C.F., Akers, N.K., Conde, L., Ladner, M., Hawbecker, S.K., Cohen, F., Ribas, F., Erlich, H.A., Goodridge, D., Trachtenberg, E.A., et al. (2012). Multi-locus HLA class I and II allele and haplotype associations with follicular lymphoma. Tissue Antigens *79*, 279–286.
18. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.
19. Foo, J.N., Smedby, K.E., Akers, N.K., Berglund, M., Irwan, I.D., Jia, X., Li, Y., Conde, L., Darabi, H., Bracci, P.M., et al. (2013). Coding variants at hexa-allelic amino acid 13 of HLA-DRB1

explain independent SNP associations with follicular lymphoma risk. Am. J. Hum. Genet. *93*, 167–172.

20. Stern, L.J., Brown, J.H., Jardetzky, T.S., Gorga, J.C., Urban, R.G., Strominger, J.L., and Wiley, D.C. (1994). Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. Nature *368*, 215–221.
21. Rapin, N., Hoof, I., Lund, O., and Nielsen, M. (2010). The MHC motif viewer: a visualization tool for MHC binding motifs. Curr. Protoc. Immunol. Unit *18*, 17.
22. Díaz, G., Amicosante, M., Jaraquemada, D., Butler, R.H., Guillén, M.V., Sánchez, M., Nombela, C., and Arroyo, J. (2003). Functional analysis of HLA-DP polymorphism: a crucial role for DPbeta residues 9, 11, 35, 55, 56, 69 and 84-87 in T cell allorecognition and peptide binding. Int. Immunol. *15*, 565–576.
23. Taylor, G.M., Gokhale, D.A., Crowther, D., Woll, P.J., Harris, M., Ryder, D., Ayres, M., and Radford, J.A. (1999). Further investigation of the role of HLA-DPB1 in adult Hodgkin's disease (HD) suggests an influence on susceptibility to different HD subtypes. Br. J. Cancer *80*, 1405–1411.
24. Conde, L., Bracci, P.M., Richardson, R., Montgomery, S.B., and Skibola, C.F. (2013). Integrating GWAS and expression data for functional characterization of disease-associated SNPs: an application to follicular lymphoma. Am. J. Hum. Genet. *92*, 126–130.
25. Sillé, F.C., Conde, L., Zhang, J., Akers, N.K., Sanchez, S., Maltbaek, J., Riby, J.E., Smith, M.T., and Skibola, C.F. (2014). Follicular lymphoma-protective HLA class II variants correlate with increased HLA-DQB1 protein expression. Genes Immun. *15*, 133–136.

The American Journal of Human Genetics, Volume *95*

Supplemental Data

Genome-wide Association Study Identifies Five Susceptibility Loci for Follicular Lymphoma outside the HLA Region

Christine F. Skibola, Sonja I. Berndt, Joseph Vijai, Lucia Conde, Zhaoming Wang, Meredith Yeager, Paul I.W. de Bakker, Brenda M. Birmann, Claire M. Vajdic, Jia-Nee Foo, Paige M. Bracci, Roel C.H. Vermeulen, Susan L. Slager, Silvia de Sanjose, Sophia S. Wang, Martha S. Linet, Gilles Salles, Qing Lan, Gianluca Severi, Henrik Hjalgrim, Tracy Lightfoot, Mads Melbye, Jian Gu, Hervé Ghesquières, Brian K. Link, Lindsay M. Morton, Elizabeth A. Holly, Alex Smith, Lesley F. Tinker, Lauren R. Teras, Anne Kricker, Nikolaus Becker, Mark P. Purdue, John J. Spinelli, Yawei Zhang, Graham G. Giles, Paolo Vineis, Alain Monnereau, Kimberly A. Bertrand, Demetrius Albanes, Anne Zeleniuch-Jacquotte, Attilio Gabbas, Charles C. Chung, Laurie Burdett, Amy Hutchinson, Charles Lawrence, Rebecca Montalvan, Liming Liang, Jinyan Huang, Baoshan Ma, Jianjun Liu, Hans-Olov Adami, Bengt Glimelius, Yuanqing Ye, Grzegorz S. Nowakowski, Ahmet Dogan, Carrie A. Thompson, Thomas M. Habermann, Anne J. Novak, Mark Liebow, Thomas E. Witzig, George J. Weiner, Maryjean Schenk, Patricia Hartge, Anneclaire J. De Roos, Wendy Cozen, Degui Zhi, Nicholas K. Akers, Jacques Riby, Martyn T. Smith, Mortimer Lacher, Danylo J. Villano, Ann Maria, Eve Roman, Eleanor Kane, Rebecca D. Jackson, Kari E. North, W. Ryan Diver, Jenny Turner, Bruce K. Armstrong, Yolanda Benavente, Paolo Boffetta, Paul Brennan, Lenka Foretova, Marc Maynadie, Anthony Staines, James McKay, Angela R. Brooks-Wilson, Tongzhang Zheng, Theodore R. Holford, Saioa Chamosa, Rudolph Kaaks, Rachel S. Kelly, Bodil Ohlsson, Ruth C. Travis, Elisabete Weiderpass, Jacqueline Clavel, Edward Giovannucci, Peter Kraft, Jarmo Virtamo, Patrizio Mazza, Pierluigi Cocco, Maria Grazia Ennas, Brian C.H. Chiu, Joseph F. Fraumeni, Jr., Alexandra Nieters, Kenneth Offit, Xifeng Wu, James R. Cerhan, Karin E. Smedby, Stephen J. Chanock, and Nathaniel Rothman

1. SUPPLEMENTAL FIGURES

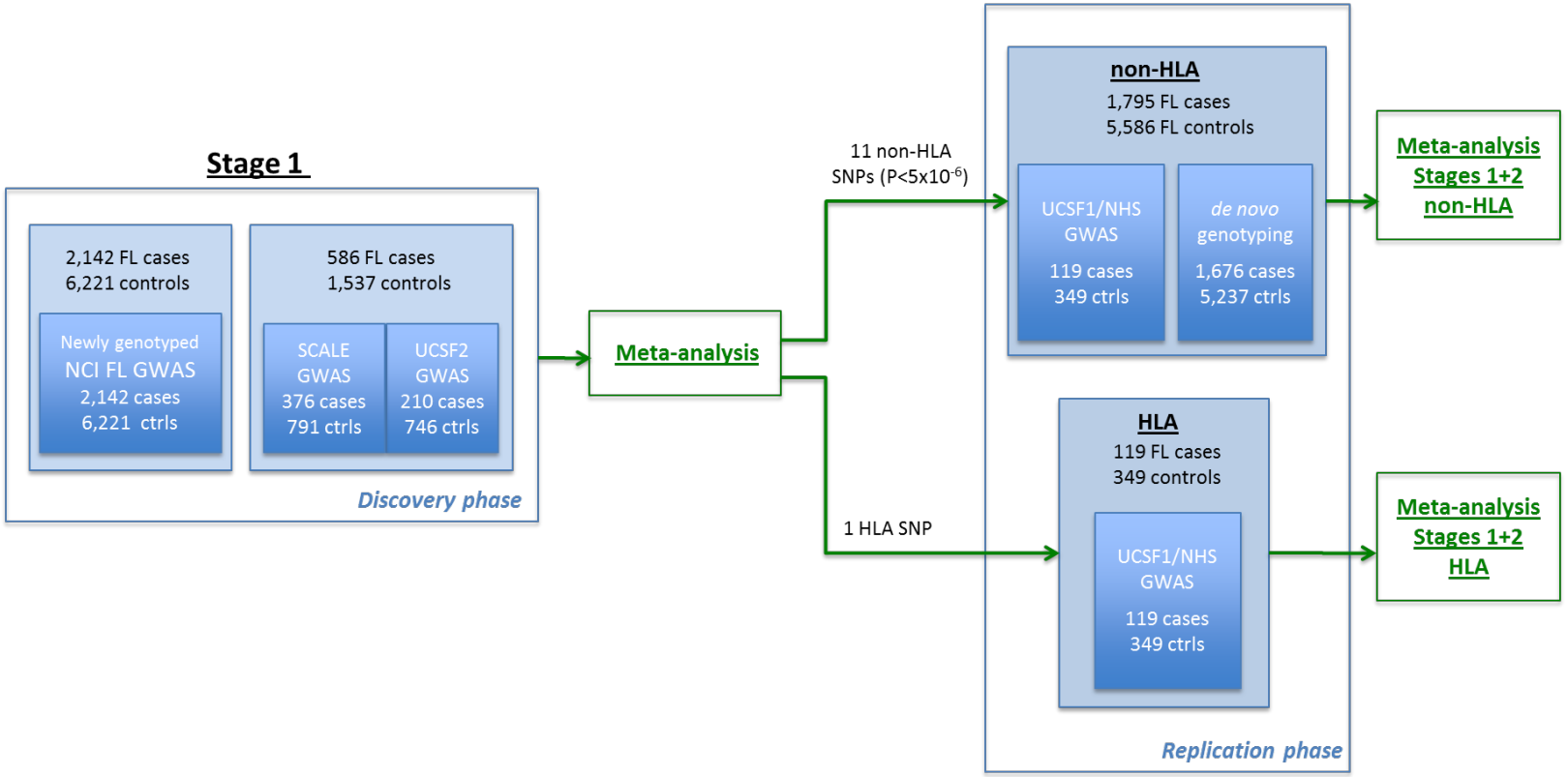


Figure S1. Schematic of the study design.

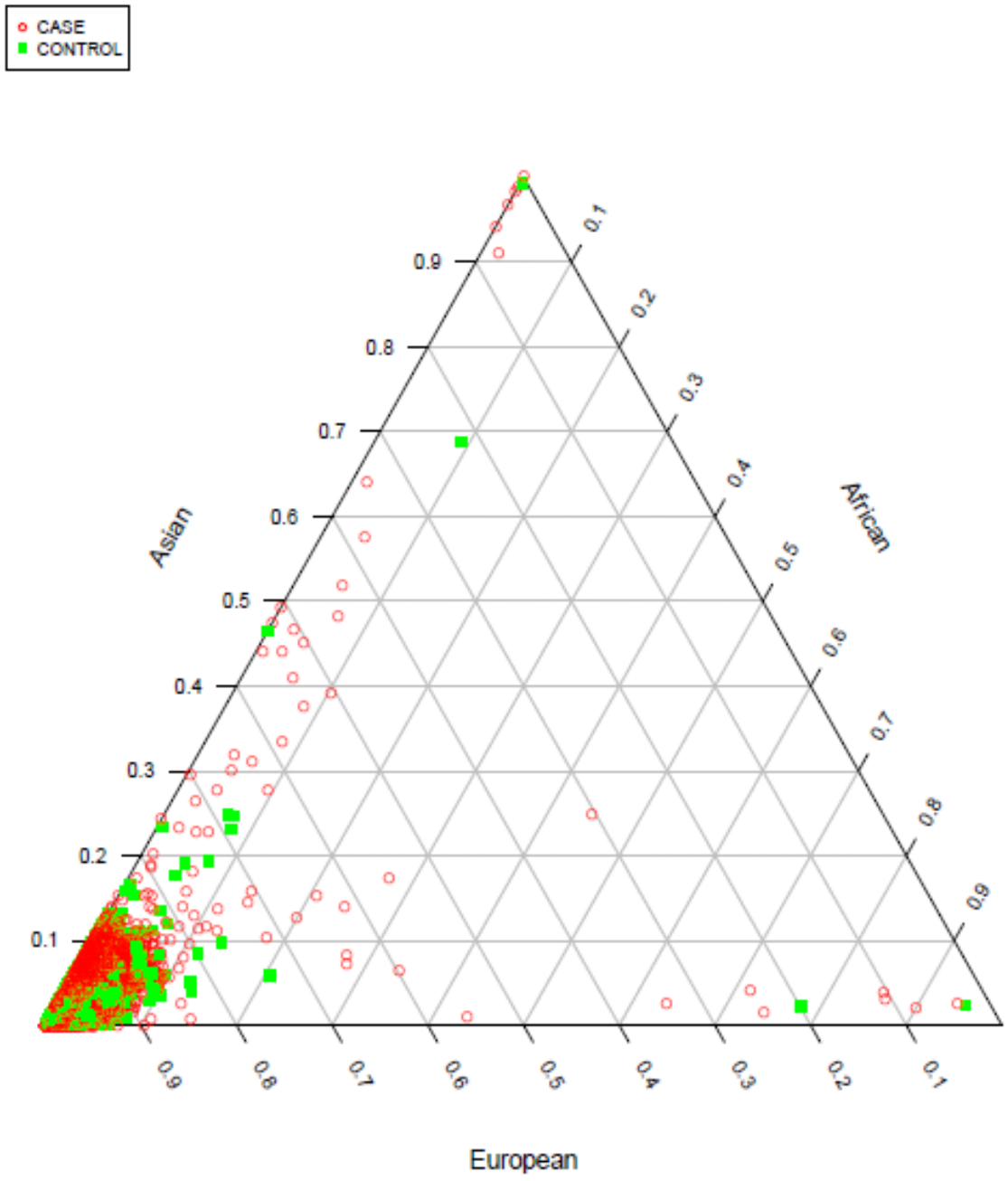


Figure S2. Plot of estimated admixture for individuals genotyped in the NCI FL GWAS. Individuals with <80% European ancestry were excluded.

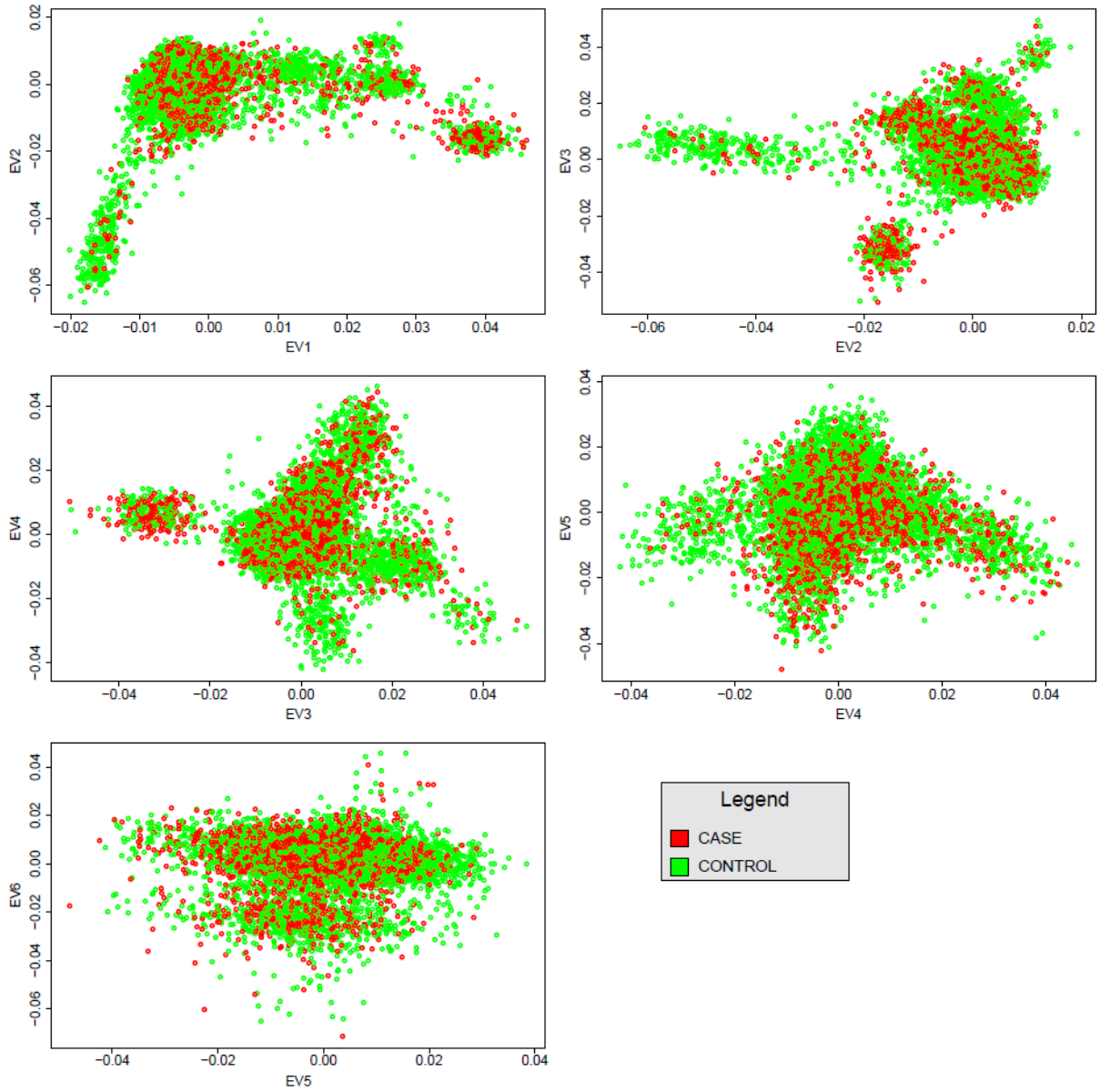


Figure S3: Plot of top eigenvectors from the NCI FL GWAS based on principal components analysis.

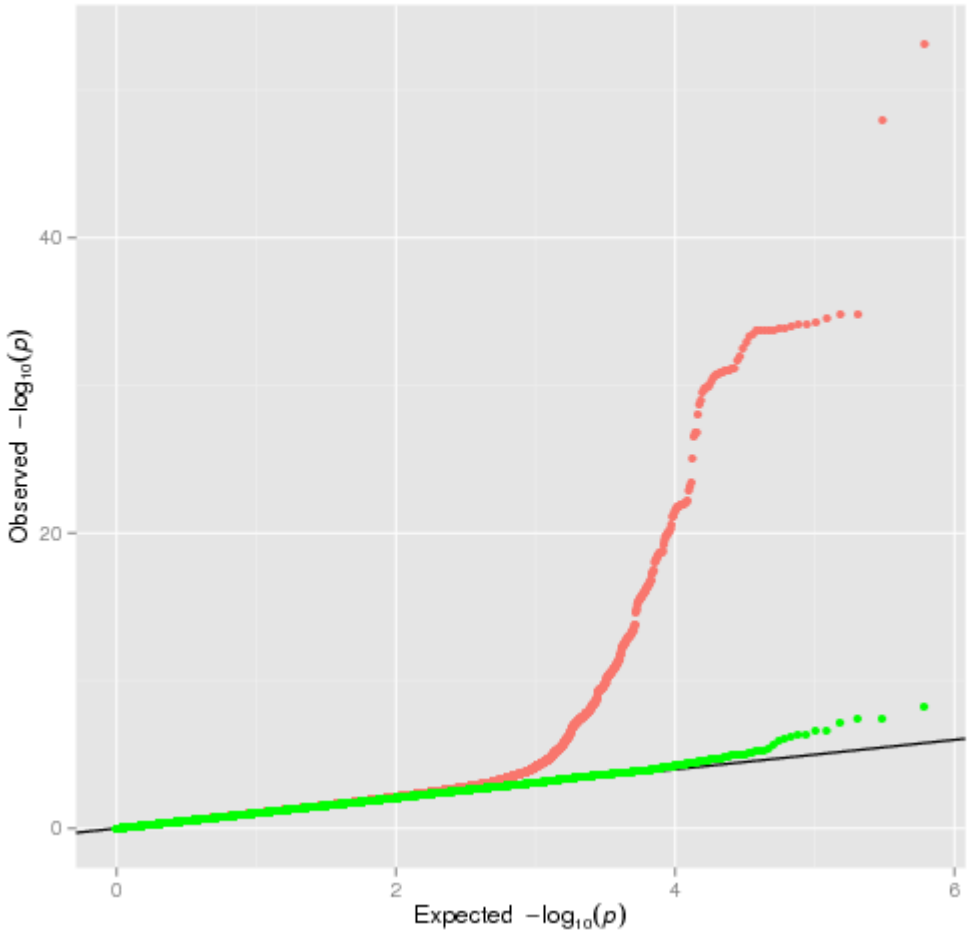


Figure S4. Quantile-quantile (Q-Q) plot of the association results for the genotyped SNPs for follicular lymphoma from the NCI FL GWAS before (red) and after removing SNPs located in the HLA region (green).

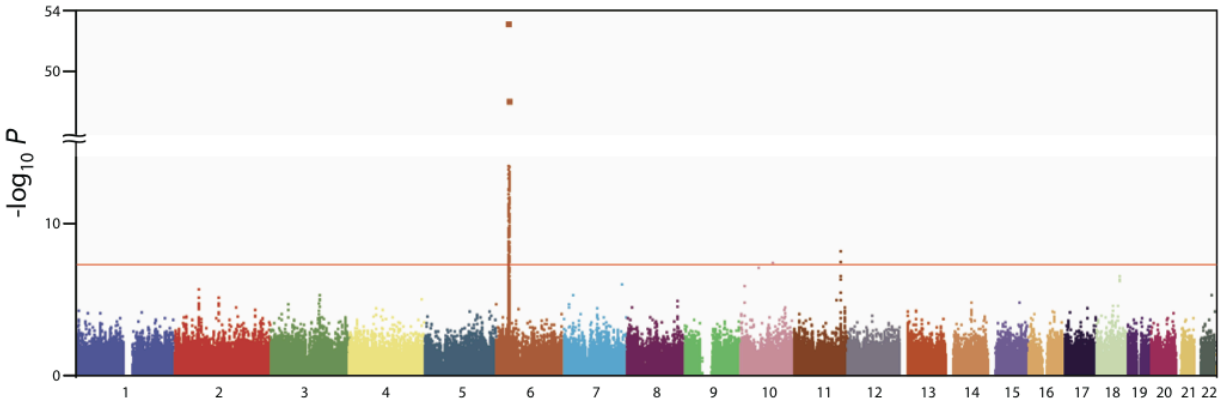
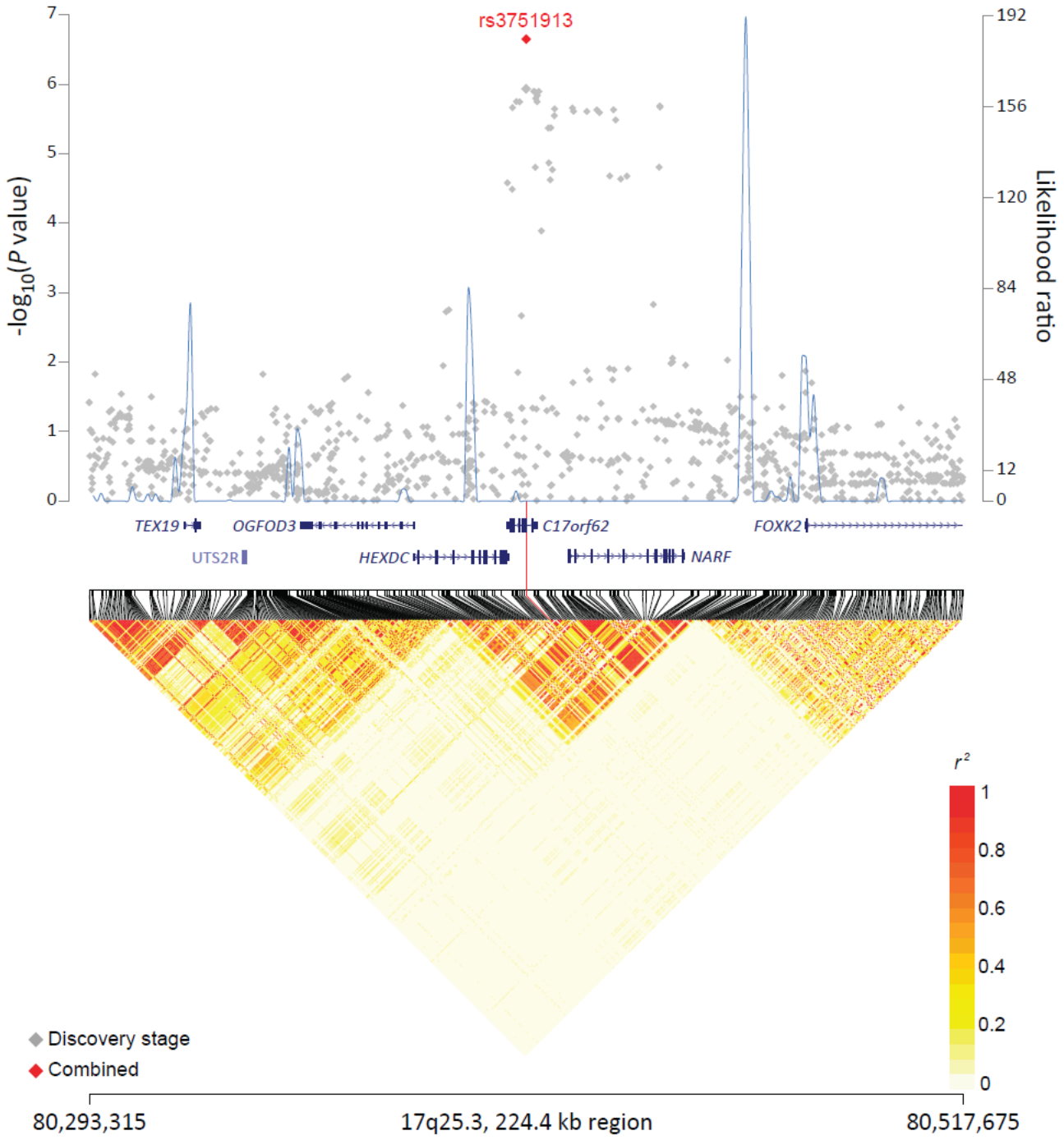
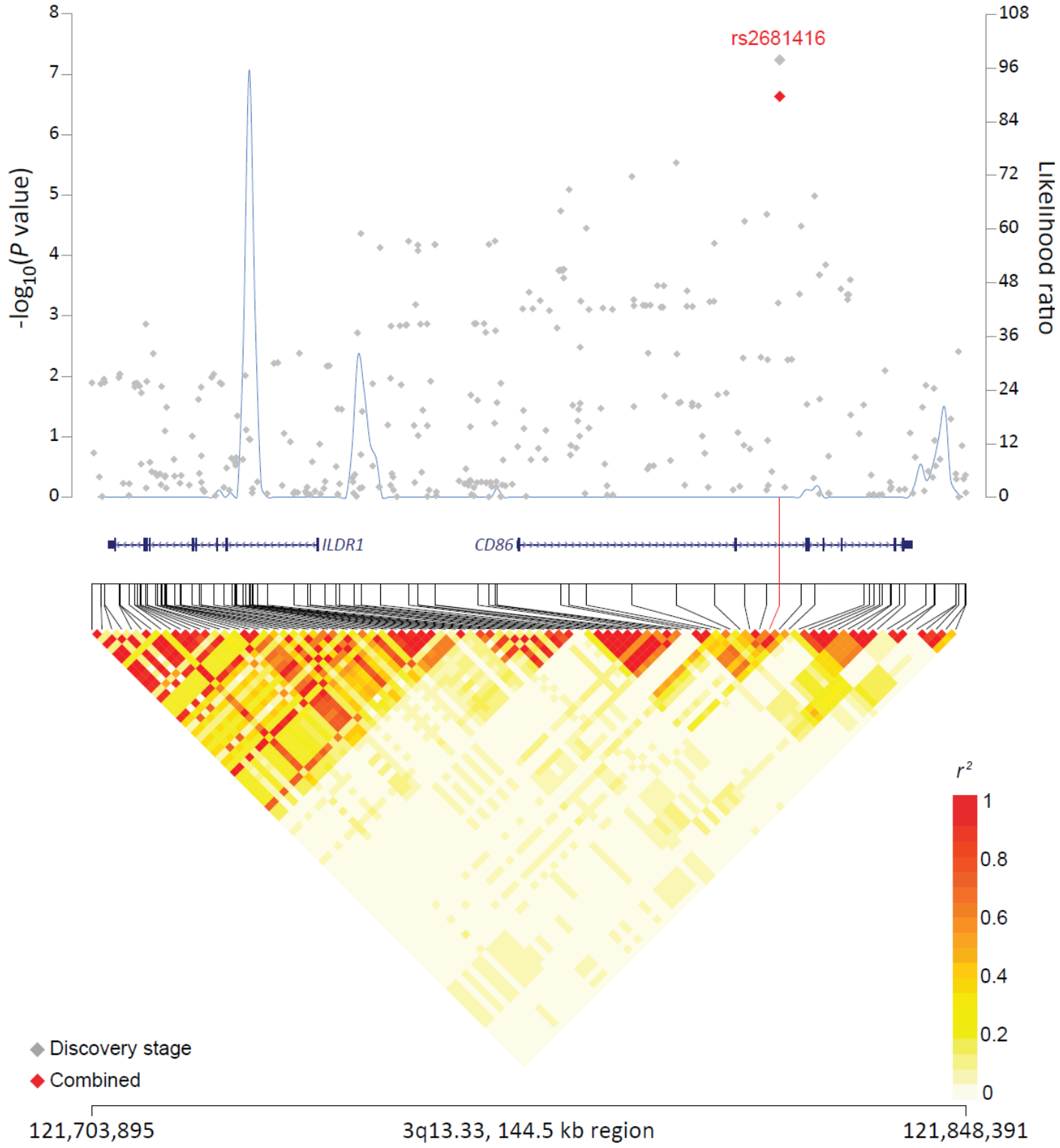


Figure S5. Genome-wide *P*-values for the association results from the NCI FL GWAS plotted against their respective chromosomal positions. Shown are the two-sided P-values obtained using the Cochran–Armitage trend test from 611,844 autosomal SNPs in 2,142 cases and 6,221 controls. The red horizontal line represents the genome-wide significance threshold level ($P<5.0\times10^{-8}$).

A.


rs3751913
-log10(P value)
Likelihood ratio
7
6
5
4
3
2
1
0
192
156
120
84
48
12
0
TEX19
OGFOD3
C17orf62
FOXK2
UTS2R
HEXDC
NARF
r²
1
0.8
0.6
0.4
0.2
0
Discovery stage
Combined
80,293,315
17q25.3, 224.4 kb region
80,517,675

B.


rs2681416
-log10(P value)
Likelihood ratio
8
7
6
5
4
3
2
1
0
108
96
84
72
60
48
36
24
12
0
ILDR1
CD86
r2
1
0.8
0.6
0.4
0.2
0
Discovery stage
Combined
121,703,895
3q13.33, 144.5 kb region
121,848,391

C.

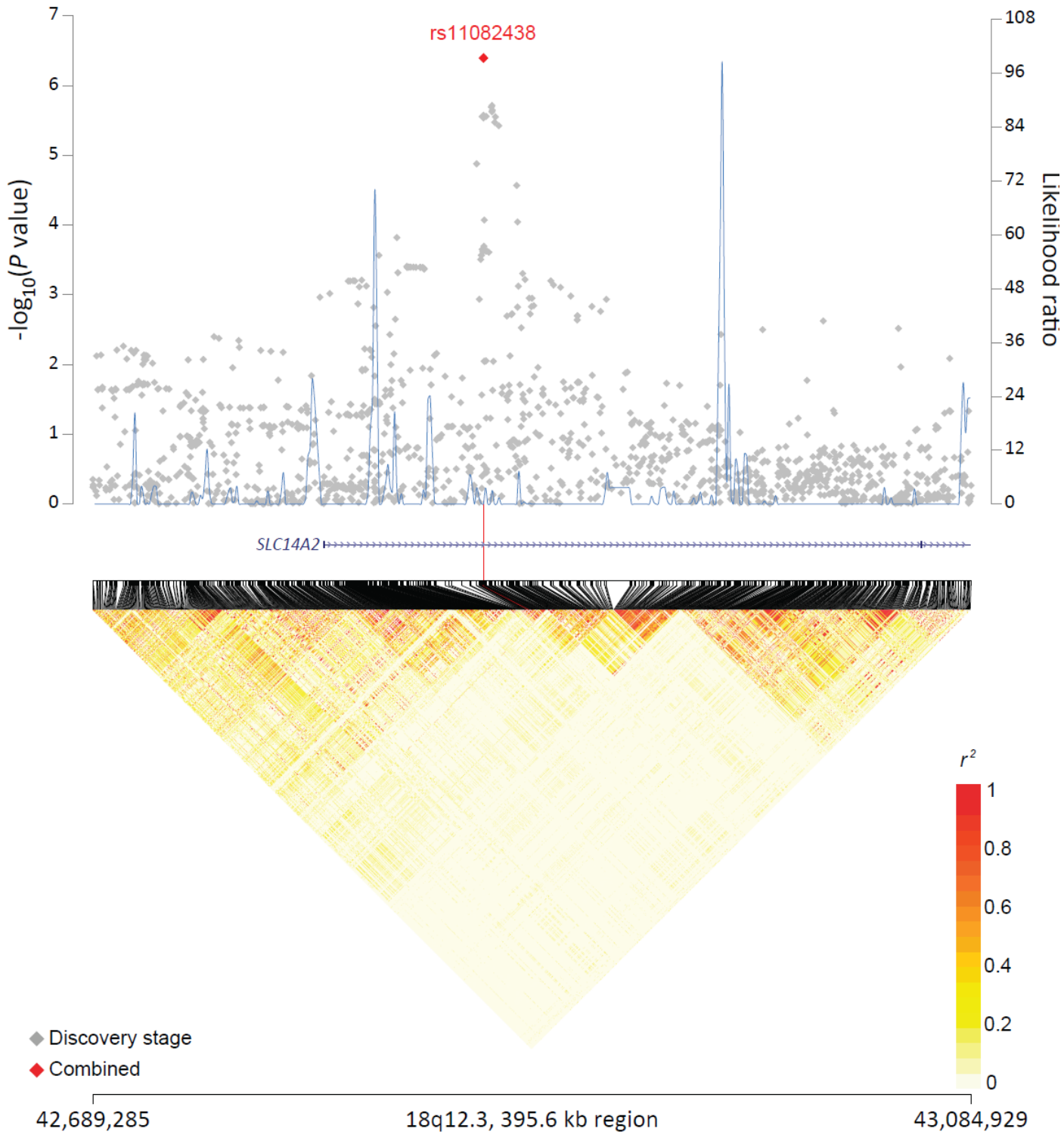


Figure S6A-C. Regional plots of the non-HLA regions approaching significance with risk of follicular lymphoma. The figure shows the association results from stage 1 (gray dots) and stages 1+2 combined (red diamond), recombination hotspots and LD plots for the associated loci: rs3751913 in 17q25.3 (A), rs2681416 in 3q13.33 (B), and rs11082438 in 18q12.3 (C).

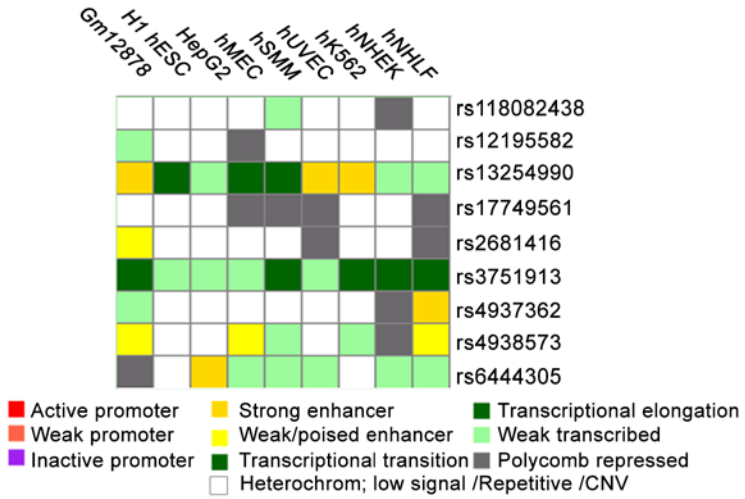
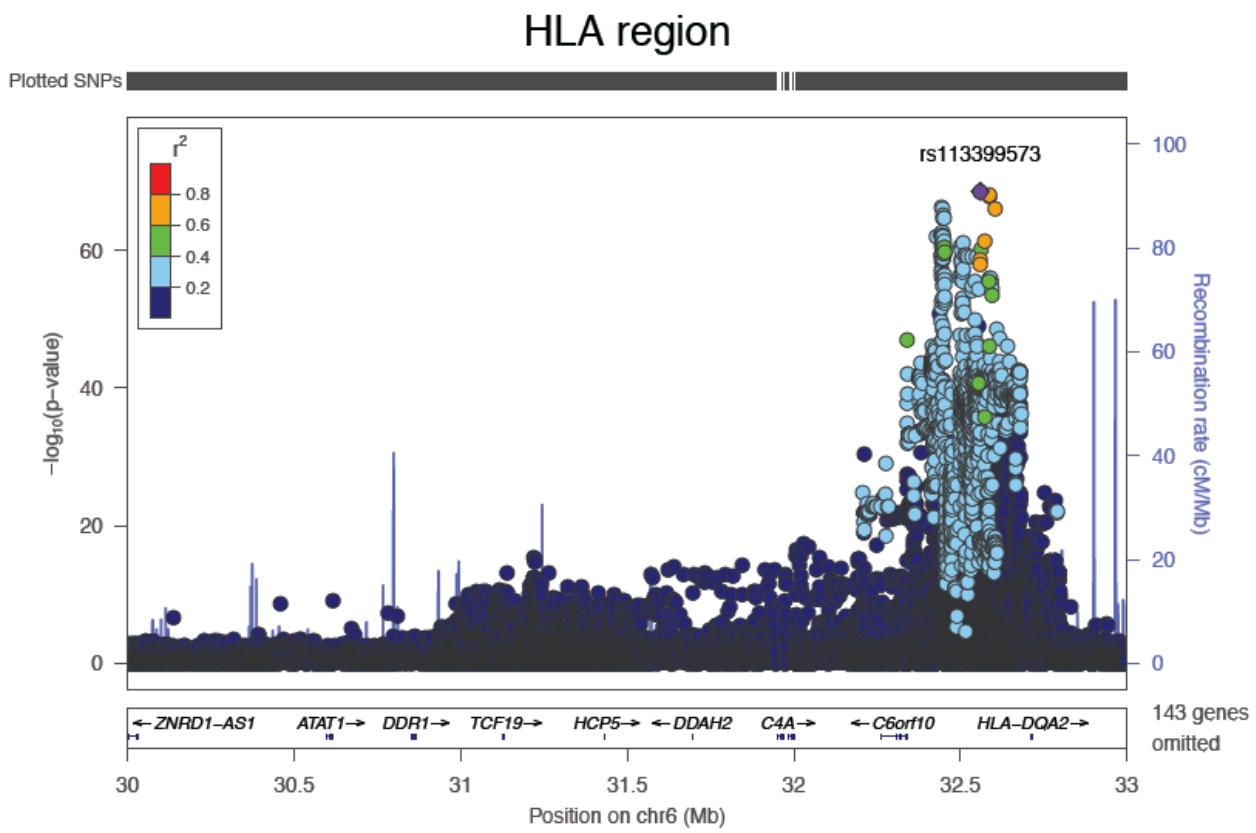


Figure S7. Chromatin states affected by follicular lymphoma associated SNPs.

A.


HLA region
Plotted SNPs
r²
0.8
0.6
0.4
0.2
rs113399573
−log10(p-value)
Recombination rate (cM/Mb)
0
20
40
60
80
100
←ZNRD1–AS1
ATAT1→
DDR1→
TCF19→
HCP5→
←DDAH2
C4A→
←C6orf10
HLA–DQA2→
143 genes omitted
30
30.5
31
31.5
32
32.5
33
Position on chr6 (Mb)

B.

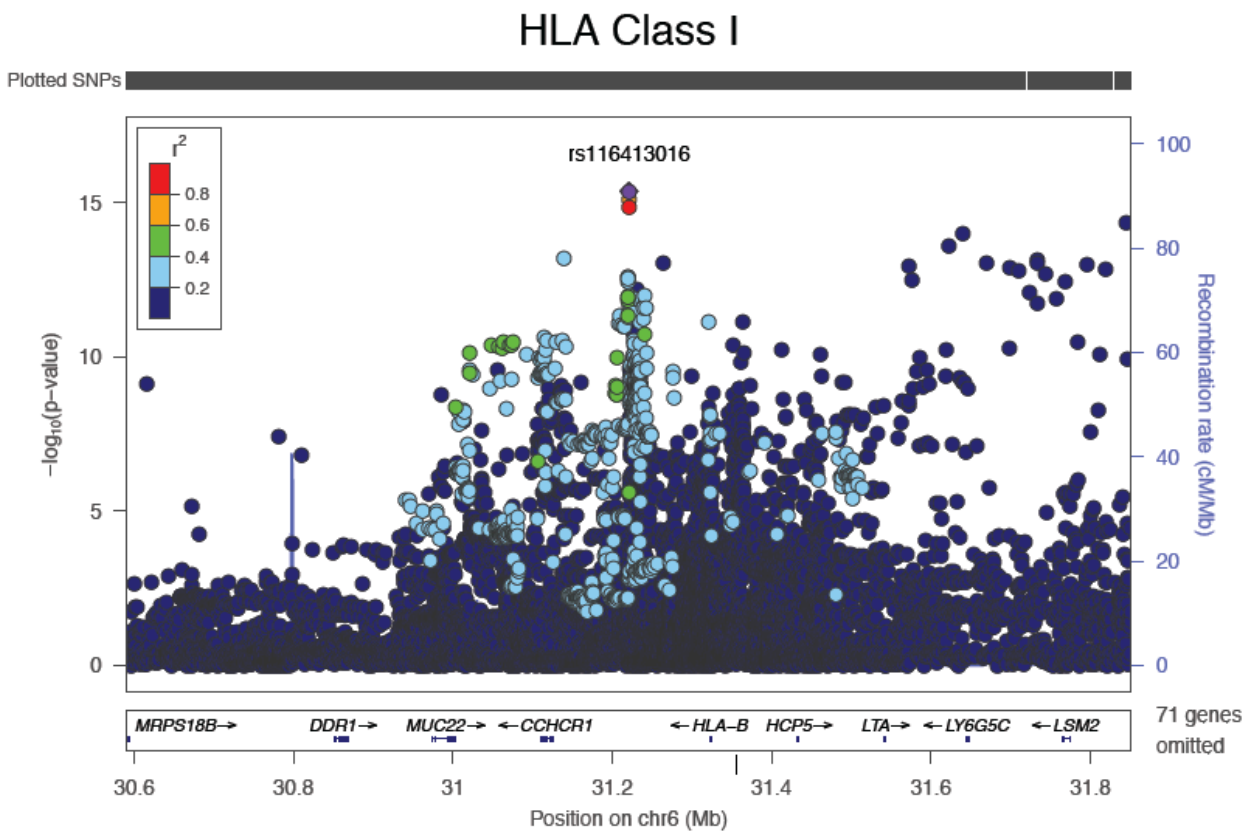


Figure S8A-B. Regional plots of the FL associated loci in the HLA region. Figure shows the association results for the HLA region (30-33Mb) (A), where the main association peak can be observed in the HLA Class II (32-33Mb), as well a zoom-in on the HLA Class I region (B), where a secondary peak can be observed approximately at 31-31.4Mb. SNPs are colored based on LD with the most significantly associated SNP in the region.

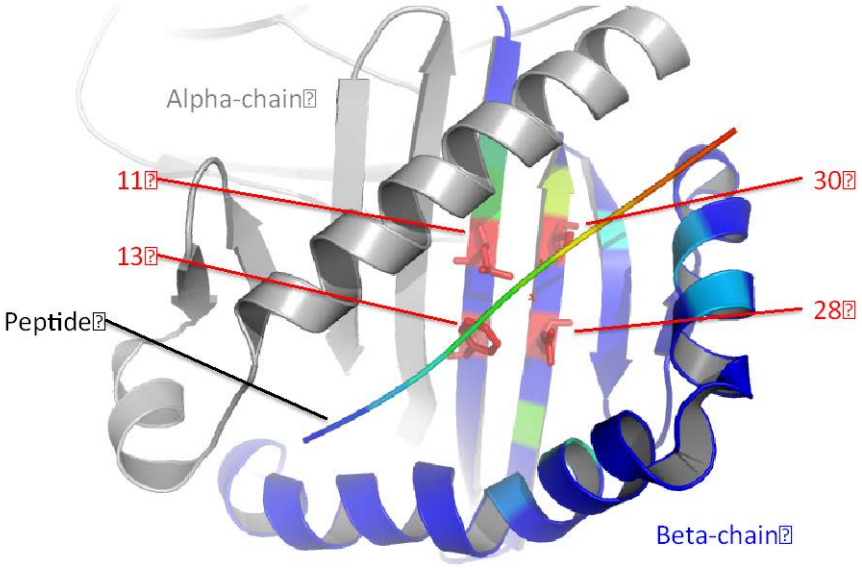


Figure S9. 3D-structural configuration of HLA-DRβ1 residues around the peptide binding groove. Amino acid residues at DRβ1 positions 11, 13, 28, and 30 are labeled and all have direct contact with the peptide in the binding groove.

2. SUPPLEMENTAL TABLES

Study Name	Study Abbreviation	No. FL Cases[a]	No. Controls[a]	Design, location	Source of cases	Source of controls
Stage 1 – NCI FL GWAS						
Cohort Studies						
Alpha-Tocopherol, Beta-Carotene Lung Cancer Prevention Study[1]	ATBC	24	240[b]	Nested case-control, Finland	Identified through linkage to the Finnish Cancer Registry	Cohort participants without a diagnosis of cancer
American Cancer Society Cancer Prevention Study-II Nutrition Cohort[2]	CPS-II	161	220[b]	Nested case-control, USA	Self-report through biannual questionnaires (starting in 1997). Verified by medical records or linkage to state cancer registry.	Cohort participants alive at time of case diagnosis without cancer
European Prospective Investigation into Cancer, Chronic Diseases, Nutrition and Lifestyles[3; 4]	EPIC	53	773	Nested case-control, multiple European countries	Cases identified through population cancer registries in seven of the participating countries (Denmark, Italy, The Netherlands, Norway, Spain, Sweden and the UK) and through a combination of methods including health insurance records, cancer and pathology registries, and by active follow-up through study subjects and their next-of-kin in three countries (France, Germany and Greece).	Cohort participants matched by age, sex and study center who were alive and cancer-free at the time of diagnosis of the corresponding case
Health Professionals Follow-up Study[5]	HPFS	5	86	Nested case-control, USA	Self-report through bi-annual questionnaires. Verified by medical records and pathology report	Cohort participants alive at time of case diagnosis without cancer, matched on date of birth, ethnicity, date and time of day of blood collection, and fasting status

Study Name	Study Abbreviation	No. FL Cases[a]	No. Controls[a]	Design, location	Source of cases	Source of controls
The Melbourne Collaborative Cohort Study[6]	MCCS	59	246	Nested case-control, Australia	Incident cases ascertained through national cancer registries	Controls were unaffected cohort participants
Nurses' Health Study[7; 8]	NHS	25	90	Nested case-control, USA	Self-report through bi-annual questionnaires. Verified by medical records and pathology report	Cohort participants alive at time of case diagnosis without cancer, matched on date of birth, ethnicity, date and time of day of blood collection, and fasting status
New York University Women's Health Study[9; 10]	NYU-WHS	13	56	Nested case-control, USA	Self-report through questionnaires every 2-4 years, confirmed by medical and pathology records; and linkages to tumor registries of NY, NJ and Florida and NDI	Cohort participants selected by incidence density sampling (alive and free of cancer at time of case diagnosis
Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial[11; 12]	PLCO	120	3076[b]	Nested case-control, USA	Self-report through annual questionnaires. Verified by medical records and pathology report	Cohort participants alive at time of case diagnosis without cancer diagnosis

Study Name	Study Abbreviation	No. FL Cases[a]	No. Controls[a]	Design, location	Source of cases	Source of controls
Women's Health Initiative[13]	WHI	176	395	Nested case-control, USA	Self-report through semi-annual clinic visits or annual contact. Verified through medical records	Cohort participants without a diagnosis of cancer
Population-based case-control studies						
British Columbia Non-Hodgkin Lymphoma Study[14]	BCCA	106	390	Population-based case-control study, Canada	First primary NHL diagnosis from Vancouver and Victoria metropolitan areas identified through the BC Cancer Registry (excluding HIV-infected and post-transplant cases)	Controls from the same areas, matched on area, age, and sex ascertained from the British Columbia Health Insurance files
Epidemiology & Genetics Unit Lymphoma Case-Control study[15-17]	ELCCS	188	461	Population-based case-control study, UK	Cases were patients aged between 18-69 residing in predefined geographic areas and newly diagnosed with NHL between 1998 and 2003. Diagnoses were pathologically confirmed and coded to the WHO Classification for Oncology	For each case, one age- and sex-matched control was randomly selected from population based General practice registers
Multicenter Italian study on gene-environment interactions in lymphoma etiology: translational aspects	Italian GxE	11	54	Population-based case-control study, Italy	First primary NHL diagnosis identified in the Hematology Departments of the participating centres	Cohort participants alive at time of case diagnosis without cancer

Study Name	Study Abbreviation	No. FL Cases[a]	No. Controls[a]	Design, location	Source of cases	Source of controls
National Cancer Institute-Surveillance, Epidemiology, and End Results Interdisciplinary Case-Control Study of Non-Hodgkin's Lymphoma[18; 19]	NCI-SEER	235	689	Population-based case-control study, USA	First primary NHL diagnosis identified through 4 SEER registries (excluding HIV-infected cases)	Controls from the same areas, matched on area, age, and race ascertained though random digit dialing (<64 years of age) and CMMS files (≥65 years of age)
NSW non-Hodgkin lymphoma study[20]	NSW	158	397	Population-based case-control study, New South Wales (NSW) and Australian Capital Territory (ACT), Australia	Incident NHL diagnosis identified through NSW or ACT cancer registry (excluding HIV-infected cases and transplant recipients)	Controls randomly selected from electoral rolls, matched on age, sex and State of residence at diagnosis
Scandinavian Lymphoma Etiology Study[21]	SCALE	0	301	Population-based case-control study, Scandinavia	Patients with incident primary NHL diagnosed through rapid case-ascertainment network in Sweden and Denmark	Frequency matched (age in 10 year intervals, sex and country) population controls prospectively identified every 6 months in nationwide population registers (incidence density sampling).
Molecular Epidemiology of non-Hodgkin lymphoma[22]	UCSF	7	10	Population-based case-control study, USA	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+

Study Name	Study Abbreviation	No. FL Cases[a]	No. Controls[a]	Design, location	Source of cases	Source of controls
Population-based case-control study in Connecticut women[23]	Yale	104	504	Population-based case-control study, USA	First primary NHL diagnosis identified through the Rapid Case Shared Resources from all the hospitals in Connecticut	Population-based controls through random digit dialing for cases <65 years and Medicare files for ≥65 years
Clinic or hospital-based or mixed case-control studies						
Environmental and genetic risks factors study in adult lymphoma[24]	ENGELA	34	278	Hospital-based case-control study, France	Recent diagnosis of a NHL as per the WHO classification (ICD-O-3) / Cases with AIDS or on immunosuppressant drugs were not eligible. Path reports for 100%, slides review for selected NHL	Hospitalized in the same hospitals as the cases, for any reason except cancer, an accident or a disease directly related to the subject's occupation, smoking, or alcohol consumption. HIV negative.
Epilymph case-control study in six European countries[25]	EpiLymph	136	1172	Multicenter case-control study, hospital-based and population-based, Europe	First primary lymphoma diagnosis (according to the 2001 WHO classification of lymphoma	Controls from Germany and Italy were randomly selected by sampling from the general population, matched to cases on gender, 5-year age-group, and residence area. The rest of the centers used matched hospital controls, with eligibility criteria limited to diagnoses other than cancer, infectious diseases and immunodeficient diseases
Iowa-Mayo SPORE Molecular Epidemiology Resource[26]	Iowa-Mayo SPORE	233	0	Clinic-based case registry, USA	Consecutive patients with newly diagnosed, histologically-confirmed non-Hodgkin lymphoma (excluding HIV-infected cases) who were residents of US	N/A

Study Name	Study Abbreviation	No. FL Cases[a]	No. Controls[a]	Design, location	Source of cases	Source of controls
Mayo Clinic Case-Control Study of NHL[27]	Mayo Case-Control	261	911	Clinic-based case-control study, USA	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
Memorial-Sloan Kettering Lymphoproliferative disorders Study[28]	MSKCC	206	9	Hospital-based case-study and NYCP controls, USA	Hospital clinic based ascertainment in a tertiary referral center	NYCP controls from same geographic area
Stage 1 - Previous GWAS						
Scandinavian Lymphoma Etiology Study[21]	SCALE	580	800	Population-based case-control study, Scandinavia	Patients with incident primary NHL diagnosed through rapid case-ascertainment network in Sweden and Denmark	Frequency matched (age in 10 year intervals, sex and country) population controls prospectively identified every 6 months in nationwide population registers (incidence density sampling).

Study Name	Study Abbreviation	No. FL Cases[a]	No. Controls[a]	Design, location	Source of cases	Source of controls
Molecular Epidemiology of non-Hodgkin lymphoma[22]	UCSF2	215	753	Population-based case-control study, USA	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+
Stage 2 - Replication studies						
Iowa-Mayo SPORE Molecular Epidemiology Resource[26]	Iowa-Mayo SPORE	105	0[c]	Clinic-based case registry, USA	Consecutive patients with newly diagnosed, histologically-confirmed non-Hodgkin lymphoma (excluding HIV-infected cases) who were residents of US	N/A
Mayo Clinic Case-Control Study of NHL[27]	Mayo Case-Control	145	388	Clinic-based case-control study, USA	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
MD Anderson lymphoma case-control study	MD Anderson	541	542	Case-control, USA	MD Anderson Cancer Center	Kelsey Seybold Clinics
Memorial-Sloan Kettering Lymphoproliferative disorders Study[29]	MSKCC	403	376	Hospital-based case-study and NYCP controls, USA	Hospital clinic based ascertainment in a tertiary referral center	NYCP controls from same geographic area
NCI Replication Study[1-28; 30]	NCI Rep	605	4296	Mixed study of population, trial and hospital-based cases and controls	Follicular cases from the stage 1 studies that did not have sufficient DNA for scanning or failed in scanning due to low completion. Also includes 402 FL cases from a randomized clinical trial of advanced follicular lymphoma evaluating maintenance therapy with rituximab after induction of response with chemotherapy plus rituximab in comparison without maintenance therapy (PRIMA)	Controls from the stage 1 studies that were not scanned or failed scanning due to low completion.

Study Name	Study Abbreviation	No. FL Cases[a]	No. Controls[a]	Design, location	Source of cases	Source of controls
Molecular Epidemiology of non-Hodgkin lymphoma[22; 31; 32]	UCSF1/NHS	120	349	Case series with population controls	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	Participants from the Nurses' Health Study that were selected for a GWAS of endometrial cancer.

[a]Number of cases and controls with DNA available; however, not all subjects had sufficient DNA for scanning and/or Taqman genotyping. Only a subset of controls with DNA was selected for scanning in stage 1.
[b]Controls scanned previously on the Illumina Omni2.5.
[c]No controls were ascertained for this study. For the replication, the Iowa-Mayo SPORE and the Mayo Case-Control studies were considered to be a single study.

Table S1. Description and study design of studies included in stages 1 and 2.

This GWAS of FL was part of a larger initiative that included participants of European descent from 22 NHL studies including 9 prospective cohort studies, and 8 population-based and 5 clinic- or hospital-based case-control studies. All studies obtained informed consent from participants and approval from the respective Institutional Review Boards for this study. Cases were ascertained from cancer registries, clinics or hospitals or through self-report verified by medical and pathology reports. The phenotype information for all NHL cases was reviewed centrally at the International Lymphoma Epidemiology Consortium (InterLymph) Data Coordinating Center and harmonized according to the hierarchical classification proposed by the InterLymph Pathology Working Group based on the World Health Organization (WHO) classification (2008)[33; 34]

Study	Sample QC				Genotyping and Imputation							
		Inclusion/exclusion criteria					Inclusion criteria for Imputation					
	No. of cases /controls in file	Minimum sample call rate for inclusion	Exclusions	No.cases /controls after exclusions	Platform	Genotype calling algorithm	MAF	SNP Call rate*	P for HWE	SNPs that met QC criteria	Imputation Software	SNPs in meta-analysis
Stage 1: Discovery												
NCI	2301/6390[a]	>=93%	1) Abnormal heterozygosity; 2) gender discordance; 3) unexpected duplicates; 4) Non CEU	2142/6221	Illumina OmniExpress	BeadStudio (GenCall)	>= 0.01	>=0.95	>=1e-6	611844	IMPUTE2	21554489[b]
SCALE	379/791	>=95%	Unexpected duplicates	376/791	Illumina HumanHap 317K	BeadStudio (GenCall)	>= 0.01	>=0.95	>=1e-6	298045	IMPUTE2	14205416
UCSF2	213/751	>=95%	1) Abnormal heterozygosity; 2) PCA outlier	210/746	Illumina HumanCNV370-Duo	BeadStudio (GenCall)	>= 0.01	>=0.95	>=1e-6	290523	IMPUTE2	15085856
Stage 2: Replication												
UCSF1/NHS	120/349	>=95%	Incomplete phenotype	119/349	Illumina OmniExpress	BeadStudio (GenCall)	>= 0.01	>=0.95	>=1e-6	614320	IMPUTE2	12776022

[a]A total of 3536 control subjects from ATBC, CPSII and PLCO cohorts that previously genotyped on Illumina Omni 2.5M chips were pooled into the NCI set.

[b]Use both MAF>0.0001 and INFO>0.3 for post imputation SNP filtering. NCI set resulted in more SNPs being retained for analysis.

Table S2. Information on genotyping methods, quality control, imputation, and analysis for GWAS.

	Genotyped Subjects		Exclusions										Previously scanned		Final subjects included in the analysis		
			High missing rate		Heterozygosity		Gender discordance		Unexpected duplicates		Non CEU						
Study	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Total**
Cohort studies																	
ATBC	20		3											240	17	240	257
CPS-II	161		16				1		2		1			220	141	220	361
EPIC	47	275	1	9				1							46	265	311
HPFS	5	86		0								1			5	85	90
MCCS	59	76	1	1											58	75	133
NHS	25	90	1	2											24	88	112
NYU-WHS	13	56	1	3					1						11	53	64
PLCO	120		4						1					3076	115	3076	3191
WHI	176	250	13	19					1		1	3			161	228	389
Subtotal	626	833	40	34			1	1	5		2	4		3536	578	4330	4908
Population-based case-control studies																	
BCCA	106	110	7	1			1								98	109	207
ELCCS	188	251	6	5				1							182	245	427
Italian GxE	16	54		8				1							16	45	61
NCI-SEER	234	298	14	23	1		1	2			1	3			217	270	487
NSW	152	157	3	2			2	1			1				146	154	300
SCALE		299		6								2				291	291
UCSF	7	10		0											7	10	17
YALE	102	149	4	3											98	146	244
Subtotal	805	1328	34	48	1		4	5			2	5			764	1270	2034
Clinic or hospital-based or mixed case-control studies																	
ENGELA	34	77	4	14											30	63	93
EpiLymph	136	250	12	35						1	1	3			123	211	334
Iowa-Mayo SPORE	233		2	0							3				228		228
Mayo Case-Control	261	357	12	12					2	1	2	1			245	343	588
MSKCC	206	9	25	5			1				6				174	4	178
Subtotal	870	693	55	66			1		2	2	12	4			800	621	1421
Grand total	**2301**	**2854**	**129**	**148**	**1**		**6**	**6**	**7**	**2**	**16**	**13**		**3536**	**2142**	**6221**	**8363**

Table S3. Subjects genotyped, quality control exclusions, and final subjects included in the NCI FL GWAS

	No. of subjects		% Male		Mean age (SD)	
Study	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**
STAGE 1: NCI FL GWAS						
ATBC	17	240	100.00%	100.00%	67.59 (6.00)	68.35 (7.67)
BCCA	98	109	44.90%	55.96%	58.68 (11.50)	60.70 (12.66)
CPSII	141	220	58.87%	49.55%	71.10 (6.75)	68.41 (6.28)
EPIC	46	265	39.13%	45.28%	60.48 (8.46)	62.65 (8.42)
Engela	30	63	53.33%	65.08%	51.83 (11.31)	55.25 (11.34)
Epilymph	123	211	44.72%	54.03%	57.19 (11.31)	59.31 (12.96)
HPFS	5	85	100.00%	100.00%	61.80 (6.82)	70.56 (8.43)
Iowa-Mayo SPORE	228	0	63.16%	NA	59.79(12.68)	NA
Italian GxE	16	45	37.50%	62.22%	52.44 (12.90)	55.38 (11.68)
Mayo case-control	245	343	52.24%	61.22%	59.93 (13.37)	60.90 (13.35)
MCCS	58	75	39.66%	52.00%	64.72 (10.86)	70.96 (7.85)
MSKCC	174	4	54.60%	0.00%	57.46 (12.19)	41.25 (9.36)
NCI-SEER	217	270	47.47%	54.07%	55.30 (12.07)	56.57 (11.94)
NHS	24	88	0.00%	0.00%	63.67 (5.80)	64.01 (6.91)
NSW	146	154	60.96%	60.39%	54.95 (9.93)	57.19 (11.13)
NYU-WHS	11	53	0.00%	0.00%	64.36 (9.40)	76.17 (9.20)
PLCO	115	3076	55.65%	95.68%	69.03 (6.73)	69.54 (6.19)
SCALE	0	291	0.00%	57.73%	NA	60.01(12.12)
UCSF	7	10	42.86%	60.00%	62.14 (10.67)	43.40 (14.83)
ELCCS	182	245	47.80%	50.61%	53.61 (7.33)	53.19 (8.19)
WHI	161	228	0.00%	0.00%	70.53 (6.41)	77.66 (6.58)
YALE	98	146	0.00%	0.00%	60.30 (12.33)	61.68 (13.47)
Total	2142	6221	45.75%	72.77%	60.30 (12.08)	66.04 (10.44)
STAGE 1: PREVIOUS GWAS						
SCALE	376	791	49.47%	31.73%	57.06 (9.47)	50.77 (11.56)
UCSF2	210	746	52.38%	57.64%	60.83 (11.88)	61.44 (13.00)
STAGE 2: REPLICATION						
UCSF1/NHS	119	349	51.26%	0.00%	58.81 (12.87)	62.52 (8.60)
MD Anderson	541	542	52.50%	52.40%	56.70 (10.87)	59.74 (10.48)
MSKCC	375	376	48.53%	18.09%	56.08 (12.53)	57.72 (11.77)
MAYO	250	385	52.80%	57.66%	60.58 (12.44)	62.32 (11.85)
NCI Replication	510	3934	50.59%	44.81%	57.82(12.11)	60.67(14.52)

Table S4. Characteristics of the cases and controls included in the final analysis for stages 1-2.

SNP	SNP location	SNP location with respect to nearest gene(s)	Observations
rs4938573	11q23.3	12.6kb upstream of *CXCR5*	• *CXCR5* augments B-cell migration[35] and activation during B-cell receptor signaling[36] • A previous study reported a suggestive association between FL risk and a weakly correlated SNP in the *CXCR5* region, rs1790192 (r^2=0.25)[37]. In our study, mutual adjustment for both loci showed that rs4938573 is independently associated with FL risk (adjusted OR=1.30, P=4.7x10^{-12}).
rs4937362	11q24.3	35kb upstream of *ETS1*	• ETS1 is a transcription factor and member of the ETS family. It is a negative regulator of Th17 and B-cell differentiation, is expressed in lymphoid cells and regulates the expression of several genes involved in lymphocyte survival and proliferation[38].
rs6444305	3q28	Overlapping *LPP,* 836.4 kb upstream of *BCL6*	• *LPP* encodes a LIM domain-containing protein and participates in cell adhesion, cell migration, proliferation and transcription dynamics[39]. • rs6444305 is located 650.5kb from rs6773854 (r^2=0.04), a SNP in the *BCL6/LPP* region previously linked with risk of B-cell lymphoma and diffuse large B-cell lymphoma (DLBCL) in East Asian populations[40]. In the current study, rs6773854 was not strongly associated with FL (OR=1.11, P=0.01).
rs17749561	18q21.33	7.4kb downstream of *BCL2*	• *BCL2* is an anti-apoptotic oncogene, and a partner gene in the hallmark FL translocation t(14;18)(q32:q21) associated with *BCL2* over-expression[41]. • rs17749561 is located near the major translocation breakpoint within the 3'UTR and it is in high linkage disequilibrium (LD) with rs4987855 (r^2=0.86), a chronic lymphocytic leukemia (CLL) risk allele[42].
rs13254990	8q24.21	Intronic to *PVT1*	• *PVT1* is a frequent site of translocations for Burkitt lymphoma, plasmacytomas[43] and aggressive B-cell lymphomas[44]. • Whereas rs13254990 is weakly correlated ($r^2 \leq 0.2$) with other nearby SNPs associated with Hodgkin lymphoma[45], CLL[46] , and numerous solid cancers[47-51], it is in high LD with a DLBCL risk locus, rs13255292 (r^2=0.93) (*manuscript submitted*), suggesting a common susceptibility locus for FL and DLBCL.

Table S5. Characteristics of the non-HLA loci that showed genome-wide significance.

chr	pos (hg19)	LD (r^2)	LD (D')	variant	Ref	Alt	EUR freq	GERP cons	SiPhy cons	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	eQTL tissues	Motifs changed	GENCODE genes	dbSNP func annot
11	118725660	0.86	0.94	rs4936441	C	G	0.83								Evi-1, Nanog, PEBP	19kb 3' of Y_RNA	
11	118726753	0.87	-0.94	rs11217066	T	G	0.17								4 altered motifs	20kb 3' of Y_RNA	
11	118726843	0.87	-0.94	rs10892299	C	T	0.17									20kb 3' of Y_RNA	
11	118729456	0.89	0.95	rs10790269	C	T	0.82								Gfi1,HDAC2,Irf	23kb 3' of Y_RNA	
11	118734000	0.9	-0.96	rs11217074	T	C	0.17								4 altered motifs	20kb 5' of CXCR5	
11	118738298	0.96	0.99	rs7119044	G	C	0.83								8 altered motifs	16kb 5' of CXCR5	
11	118739056	0.95	0.98	rs7122669	A	G	0.82									15kb 5' of CXCR5	
11	118740104	0.93	0.98	rs201150316	AT	A	0.83								21 altered motifs	14kb 5' of CXCR5	
11	118740418	0.96	0.99	rs7951740	G	C	0.83			HepG2	HMEC, NHLF		HAE2F1		Ets,Zbtb3	14kb 5' of CXCR5	
11	118740864	0.96	0.99	rs4936443	C	T	0.82			HepG2, K562, GM12878	4 cell types	4 cell types	POL2, MAX		Zfx	14kb 5' of CXCR5	
11	118740931	0.96	0.99	rs4938572	C	T	0.82			HepG2, K562, GM12878	4 cell types	60 cell types	POL2, PBX3, MAX		4 altered motifs	14kb 5' of CXCR5	

11	118741157	0.97	1	rs7117261	T	C	0.83		HepG2, GM12878, HSMM	4 cell types	HepG2	GR, POL2	EWSR1-FLI1, Ets, NF-kappaB	13kb 5' of CXCR5
11	118741842	1	1	rs4938573	C	T	0.82			GM12878, HMEC, NHLF	Th1		4 altered motifs	13kb 5' of CXCR5
11	118742724	0.83	0.99	rs7125333	T	C	0.8			GM12878			6 altered motifs	12kb 5' of CXCR5
11	118742800	0.98	-1	rs74541740	G	A	0.18			GM12878			23 altered motifs	12kb 5' of CXCR5
11	118743286	0.82	-0.99	rs12365699	G	A	0.15			GM12878	13 cell types	7 bound proteins	HNF4, NRSF, ZID	11kb 5' of CXCR5
11	118743338	0.97	0.99	rs4936444	T	C	0.82			GM12878	9 cell types	4 bound proteins		11kb 5' of CXCR5
11	118743772	0.96	0.98	rs6421571	T	C	0.82			GM12878	AG04449			11kb 5' of CXCR5
11	118744396	0.94	0.97	rs7117313	C	G	0.82						NRSF, TATA, Zfx	10kb 5' of CXCR5
11	118744701	0.83	0.95	rs138252393	A	AAAAG	0.81						EWSR1-FLI1, HDAC2, p300	9.8kb 5' of CXCR5
11	118745243	0.96	1	rs7481797	A	G	0.83						LXR,Nkx2	9.2kb 5' of CXCR5
11	118745278	0.96	1	rs7481819	C	T	0.83						4 altered motifs	9.2kb 5' of CXCR5
11	118745884	0.95	0.99	rs10790275	G	C	0.83						Znf143	8.6kb 5' of CXCR5
11	128488322	0.8	0.98	rs2156698	G	A	0.56							RP11-264E20.1

11	128489380	0.99	1	rs4245081	C	T	0.61	GM12878	iPS			HNF4, RXR::LXR	RP11-264E20.1	
11	128489535	0.92	1	rs4245082	C	T	0.59	GM12878				7 altered motifs	RP11-264E20.1	
11	128489818	0.99	1	rs4492838	C	T	0.6	GM12878	7 cell types	NFKB, GATA2		BDP1,Hsf	RP11-264E20.1	
11	128491023	0.98	0.99	rs11221396	C	T	0.61	GM12878, HSMM	GM19240			4 altered motifs	RP11-264E20.1	
11	128492401	0.99	1	rs10893897	T	C	0.6	NHLF, HSMM	6 cell types			Gfi1,RFX5	689bp 5' of RP11-264E20.1	
11	128492571	1	1	rs11221397	G	T	0.61	NHLF, HSMM				6 altered motifs	859bp 5' of RP11-264E20.1	
11	128492739	1	1	rs4937362	T	C	0.61	NHLF	LNCaP, 8988T, Adult_CD4_Th0			5 altered motifs	1kb 5' of RP11-264E20.1	
11	128494441	0.9	0.97	rs7105899	G	A	0.61	GM12878, H1	15 cell types	NFKB,CEBPB		9 altered motifs	2.7kb 5' of RP11-264E20.1	
3	188298562	0.98	0.99	rs9878956	C	G	0.72		HL-60,NH-A			9 altered motifs	LPP	intronic
3	188298688	0.98	0.99	rs9823790	T	C	0.72		HL-60			Sox	LPP	intronic
3	188298989	0.99	1	rs1849913	G	A	0.72	HepG2				4 altered motifs	LPP	intronic
3	188299423	0.99	1	rs1849911	A	C	0.72	HepG2	RPTEC	HNF4A, HNF4G, P300		Myc,Pax-5,Pax-8	LPP	intronic
3	188299793	0.99	1	rs6444303	A	G	0.72	HepG2	5 cell types				LPP	intronic

3	188299819	0.99	1	rs6444304	C	A	0.72				HepG2	4 cell types			COMP1	LPP	intronic
3	188299902	1	1	rs6444305	G	A	0.72				HepG2				PRDM1,TATA	LPP	intronic
3	188300051	0.96	1	rs13063967	C	A,T	0.71				HepG2					LPP	intronic
18	60783211	1	1	rs17749561	G	A	0.11								Irf,NF-kappaB	7.4kb 3' of BCL2	
18	60783729	1	1	rs17676919	A	G	0.11								Brachyury,FAC1,Myc	6.8kb 3' of BCL2	
18	60785638	1	1	rs17676949	G	A	0.11									4.9kb 3' of BCL2	
18	60788745	0.96	1	rs77551289	A	G	0.11								Foxp1,GATA	1.8kb 3' of BCL2	
18	60793494	0.95	0.99	rs4987856	C	T	0.11					CD34+_Mobilized, HIPEpiC			Pax-5	BCL2	3'-UTR
18	60793549	0.95	0.99	rs4987855	C	T	0.11					CD34+_Mobilized, HIPEpiC			4 altered motifs	BCL2	3'-UTR
18	60795188	0.92	0.96	rs4987845	C	T	0.11								Pax-5,RXRA	BCL2	3'-UTR
8	129076451	1	1	rs13254990	C	T	0.36				K562, Huvec, GM12878	6 cell types	USF2, NFKB, CMYC		Pax-5,Pou3f1	PVT1	intronic
8	129076573	0.92	0.98	rs13255292	C	T	0.37				K562, Huvec, GM12878	14 cell types	8 bound proteins		Arnt,SP1	PVT1	intronic
17	80402045	0.93	0.98	rs67228117	T	C	0.13				HMEC				ERalpha-a,Esr2,Rad21	C17orf62	intronic
17	80403059	0.99	1	rs60482040	C	G	0.12					LNCaP,Ishikawa	ERALPHA_A, POL2		4 altered motifs	C17orf62	intronic

17	80403894	0.98	0.99	rs2291394	G	A	0.12			HSMM			10 altered motifs	C17orf62	intronic
17	80405552	1	1	rs3751913	T	C	0.12						6 altered motifs	C17orf62	intronic
17	80407480	0.99	1	rs60529510	C	T	0.12	4 cell types	4 cell types	11 cell types	19 bound proteins		GR	C17orf62	5'-UTR
17	80407948	0.98	0.99	rs12450191	C	T	0.12	9 cell types					9 altered motifs	C17orf62	intronic
17	80407978	0.98	0.99	rs12450192	C	T	0.12	9 cell types		GM12864,HL-60			Myc	C17orf62	intronic
17	80408348	0.98	0.99	rs9911854	A	G	0.12	9 cell types		30 cell types	5 bound proteins		BHLHE40	C17orf62	intronic
17	80408419	0.98	0.99	rs9912066	A	C	0.12	9 cell types		31 cell types	8 bound proteins		NRSF,Sin3Ak-20	C17orf62	5'-UTR
17	80408815	0.96	0.99	rs9303029	C	T	0.12	8 cell types	H1	80 cell types	15 bound proteins		11 altered motifs	109bp 5' of C17orf62	
17	80410445	0.96	0.99	rs202020100	G	GACT	0.12		K562	Th1			ATF3,HNF6	1.7kb 3' of RP13-20L14.6	
17	80410447	0.96	0.99	rs145963735	C	CTAA	0.12		K562	Th1			ATF3,HNF6	1.7kb 3' of RP13-20L14.6	
17	80411207	0.92	0.99	rs9896620	C	G	0.13		K562	K562	TAL1		NF-AT,YY1	941bp 3' of RP13-20L14.6	
17	80411883	0.93	0.98	rs58441535	A	G	0.13			HCT-116,WERI-Rb-1				265bp 3' of RP13-20L14.6	
17	80412797	0.94	0.99	rs79487483	A	G	0.13						E2F,TBX5,Zfx	RP13-20L14.6	
17	80412867	0.96	0.99	rs72857495	A	C	0.12							RP13-20L14.6	

17	80417443	0.98	0.99	rs72859113	G	C	0.12	8 cell types	Huvec	K562			NARF	intronic
17	80417729	0.98	0.99	rs78720068	C	G	0.12	GM12878	Huvec, HepG2, K562			STAT,TR4	NARF	intronic
17	80421110	0.95	0.99	rs72859120	C	T	0.12		K562			Cart1,Hsf,Pdx 1	NARF	intronic
17	80423712	0.95	0.99	rs12450240	G	T	0.12						NARF	intronic
17	80424386	0.96	0.99	rs9891900	T	G	0.12					5 altered motifs	NARF	intronic
17	80427975	0.96	0.99	rs72859125	G	A	0.12					Crx,Pitx2	NARF	intronic
17	80428532	0.96	0.99	rs12600876	G	A	0.12					4 altered motifs	NARF	intronic
17	80429066	0.96	0.99	rs111533778	TC	T	0.12			HL-60,SAEC		6 altered motifs	NARF	intronic
17	80439907	0.94	0.98	rs9898723	A	G	0.12			Fibrobl		TFII-I	NARF	intronic
17	80439925	0.94	0.98	rs9905527	T	C	0.12			Fibrobl			NARF	intronic
3	121817613	1	1	rs2681416	G	A	0.28		GM12878	Th2		Zbtb12	CD86	intronic
18	42865210	1	1	rs11082438	G	T	0.05			FibroP		5 altered motifs	217kb 3' of SETBP1	intronic
18	42866471	0.94	1	rs11660485	A	C	0.04					Pou1f1	218kb 3' of SETBP1	intronic

Table S6. Results from HaploReg analysis of newly discovered and promising follicular lymphoma risk loci outside HLA and their correlated ($r^2 > 0.80$) surrogates.

AA/Allele/SNP	Study	Effect allele[a]	Other allele[a]	EAF[b]	Information[c]	OR	95% CI	*P*	P_{het}
AA DRβ1 28 Asp	NCI	P	A	0.684	0.999	0.51	(0.47-0.55)	1.54E-61	
AA DRβ1 28 Asp	UCSF2	P	A	0.676	1.003	0.68	(0.54-0.86)	0.001	
AA DRβ1 28 Asp	SCALE	P	A	0.712	0.976	0.60	(0.49-0.74)	1.34E-06	
AA DRβ1 28 Asp	UCSF1/NHS	P	A	0.683	0.990	0.42	(0.29-0.60)	2.61E-06	
AA DRβ1 28 Asp	Combined					0.53	(0.50-0.57)	6.10E-72	0.03
AA DRβ1 28 Glu	NCI	P	A	0.304	1.006	1.95	(1.80-2.11)	1.78E-60	
AA DRβ1 28 Glu	UCSF2	P	A	0.315	0.996	1.45	(1.15-1.82)	0.002	
AA DRβ1 28 Glu	SCALE	P	A	0.262	0.995	1.57	(1.27-1.94)	2.28E-05	
AA DRβ1 28 Glu	UCSF1/NHS	P	A	0.308	1.007	2.32	(1.61-3.34)	5.92E-06	
AA DRβ1 28 Glu	Combined					1.86	(1.74-2.00)	7.99E-69	0.02
HLA-DRB1*01	NCI	P	A	0.137	1.003	1.89	(1.71-2.10)	9.56E-35	
HLA-DRB1*01	UCSF2	P	A	0.144	1.025	1.68	(1.27-2.24)	0.0003	
HLA-DRB1*01	SCALE	P	A	0.125	0.992	1.71	(1.30-2.25)	0.0001	
HLA-DRB1*01	UCSF1/NHS	P	A	0.129	1.044	1.92	(1.22-3.03)	0.005	
HLA-DRB1*01	Combined					1.85	(1.70-2.03)	2.57E-42	0.81

[a]P=presence of allele, A=absence of allele
[b]EAF=effect allele frequency
[c]Information is the imputation quality score (r^2) from Beagle

Table S7. Top HLA amino acid and allele associations from the univariate analysis of HLA imputations.

AA Position	df	P_upperbound	P_lowerbound
AA DRB1 11	5	8.35e-70	5.41e-70
AA DRB1 13	5	4.11e-70	2.67e-70
AA DRB1 28	2	3.84e-67	2.49e-67
AA DRB1 30	4	4.20e-66	2.72e-66

Table S8. Global omnibus tests for associations between follicular lymphoma risk and DRβ1 AA positions at 11, 13, 28 and 30.

		Amino acid position				FL association model			
DRB1 Allele	Frequency	11	13	28	30	OR	95% CI lower	95% CI upper	*P*
*01	0.137	L	F	E	C	1.85	1.70	2.02	2.91E-42
*03:01	0.119	S	S	D	Y	0.89	0.81	0.99	2.55E-02
*04	0.16	V	H	D	Y	1.04	0.96	1.14	3.20E-01
*07:01	0.14	G	Y	E	L	1.52	1.39	1.66	1.59E-20
*08:01	0.024	S	G	D	Y	0.82	0.66	1.03	8.28E-02
*09:01	0.012	D	F	H	G	1.34	1	1.79	5.03E-02
*10:01	0.009	V	F	E	R	NA	NA	NA	NA
*11	0.09	S	S	D	Y	0.62	0.55	0.71	3.59E-13
*12:01	0.018	S	S	E	H	1.21	0.96	1.52	1.12E-01
*13	0.112	S	S	D	Y	0.56	0.50	0.63	5.32E-22
*14:01	0.022	S	S	D	Y	0.80	0.62	1.02	7.22E-02
*15:01	0.124	P	R	D	Y	0.67	0.60	0.74	1.01E-13
*16:01	0.015	P	R	D	Y	0.88	0.66	1.18	3.95E-01

Table S9. Follicular lymphoma (FL)-associated amino acid positions in HLA-DRβ1. Residue combinations of the DRβ1 amino acid positions 11, 13, 28, and 30 were associated with an increased risk of FL. Each DRB1 allele is shown with the observed frequency in controls and the amino acids carried by that allele, found in the IMmunoGeneTics/HLA Database, Release 3.15.0 (http://www.ebi.ac.uk/ipd/imgt/hla/).

LOCUS	GROUP	EFFECT ALLELE	REF ALLELE	EFFECT ALLELE FREQ	INFO	OR	95%CI lower	95%CI upper	*P*	*P* heterog.	I^2
Step 0: Unconditional association analysis											
AA DRB1 Glu-28	NCI	P	A	0.3039	1.0065	1.946	1.797	2.107	1.63E-60		
AA DRB1 Glu-28	SF	P	A	0.3149	0.996	1.45	1.152	1.824	0.00155		
AA DRB1 Glu-28	SCALE	P	A	0.2621	0.9946	1.565	1.27	1.928	2.62E-05		
AA DRB1 Glu-28	SF2	P	A	0.3087	1.0049	2.329	1.618	3.352	5.38E-06		
AA DRB1 Glu-28	NCI+SCALE+SF+SF2					1.861	1.736	1.995	7.89E-69	1.98E-02	69.56
AA DRB1 His-28	NCI	P	A	0.0122	0.9688	1.211	0.864	1.697	0.2661		
AA DRB1 His-28	SCALE	P	A	0.0253	0.9618	1.78	0.999	3.174	0.05048		
AA DRB1 His-28	NCI+SCALE					1.336	0.998	1.787	5.16E-02	2.59E-01	21.56
rs17203612	NCI	A	T	0.488	0.9648	1.903	1.759	2.059	9.39E-58		
rs17203612	SF	A	T	0.5092	0.9752	1.439	1.15	1.802	0.00148		
rs17203612	SCALE	A	T	0.5025	0.8121	1.769	1.431	2.187	1.34E-07		
rs17203612	SF2	A	T	0.4952	0.9015	1.833	1.271	2.643	0.00117		
rs17203612	NCI+SCALE+SF+SF2					1.837	1.715	1.968	4.57E-67	1.43E-01	44.79
rs3130437	NCI	A	C	0.6162	1.0059	1.288	1.192	1.393	1.72E-10		
rs3130437	SF	A	C	0.6429	1.0312	1.673	1.309	2.139	3.91E-05		
rs3130437	SCALE	A	C	0.5857	0.9768	1.296	1.069	1.57	0.00811		
rs3130437	SF2	A	C	0.6585	1.0378	1.859	1.278	2.703	0.00118		
rs3130437	NCI+SCALE+SF+SF2					1.332	1.244	1.425	1.59E-16	6.77E-02	57.96
Step 1: Conditional association analysis (conditioning on both AA DRB1 Glu-28 and AA DRB1 His-28)											
rs17203612	NCI	A	T	0.488	0.9647	1.469	1.328	1.624	6.55E-14		
rs17203612	SF	A	T	0.5092	0.9752	1.247	0.941	1.652	0.1245		
rs17203612	SCALE	A	T	0.5025	0.8121	1.504	1.166	1.939	0.00164		
rs17203612	SF2	A	T	0.4952	0.9015	1.087	0.678	1.743	0.7278		
rs17203612	NCI+SCALE+SF+SF2					1.435	1.315	1.566	4.59E-16	4.53E-01	0
rs3130437	NCI	A	C	0.6162	1.006	1.191	1.1	1.29	1.72E-05		
rs3130437	SF	A	C	0.6429	1.0312	1.675	1.308	2.145	4.45E-05		
rs3130437	SCALE	A	C	0.5857	0.9768	1.225	1.007	1.491	0.04234		
rs3130437	SF2	A	C	0.6585	1.0378	1.617	1.098	2.38	0.01487		
rs3130437	NCI+SCALE+SF+SF2					1.241	1.157	1.33	1.32E-09	3.75E-02	64.53

Step 2: Conditional association analysis (conditioning on AA DRB1 Glu-28, AA DRB1 His-28 and rs114137077)											
rs3130437	NCI	A	C	0.6162	1.006	1.179	1.088	1.277	5.82E-05		
rs3130437	SF	A	C	0.6429	1.0312	1.672	1.305	2.144	4.94E-05		
rs3130437	SCALE	A	C	0.5857	0.9768	1.208	0.992	1.472	0.06061		
rs3130437	SF2	A	C	0.6585	1.0378	1.629	1.105	2.401	0.01371		
rs3130437	NCI+SCALE+SF+SF2					1.229	1.145	1.318	8.23E-09	2.92E-02	66.7

Table S10. Results of HLA forward stepwise conditional analysis.

SNP	Position	Study	Genotyped or imputed (info)	No. of cases/ controls	Effect allele/ other allele	EAF	OR	(95%CI)	*P*	*P conditioned on DRB1 position 28, rs17203612 and rs3130437*	Reference
rs6457327	31074030	NCI	g	2142/6221	C/A	0.617	1.21	(1.12-1.31)	1.22E-06		
		UCSF2	g	210/746	C/A	0.636	1.56	(1.25-1.96)	0.0001		
		SCALE	g	376/791	C/A	0.560	1.23	(1.02-1.48)	0.03		
		UCSF1/NHS	i (0.998)	118/349	C/A	0.638	1.58	(1.12-2.24)	0.01		
		Combined		2846/8107			1.25	(1.17-1.34)	3.96E-11	0.10	(OR=1.68, $P=4.7x10^{-11}$)[52]
rs3132453	31604044	NCI	g	2142/6221	G/T	0.933	1.24	(1.06-1.44)	0.006		
		UCSF2	g	210/746	G/T	0.937	1.71	(1.05-2.78)	0.03		
		SCALE	g	376/791	G/T	0.920	1.34	(0.94-1.92)	0.11		
		UCSF1/NHS	i (0.862)	119/348	G/T	0.943	1.18	(0.56-2.47)	0.66		
		Combined		2847/8106			1.28	(1.12-1.46)	0.0003	0.26	(OR=1.69, $P=1.26\times10^{-6}$)[53]
rs9268853	3242643	NCI	g	2142/6221	C/T	0.306	1.4	(1.29-1.51)	3.03E-17		
		UCSF2	i (0.999)	210/746	C/T	0.339	1.05	(0.84-1.32)	0.66		
		SCALE	i (0.998)	376/790	C/T	0.310	1.31	(1.09-1.58)	0.005		
		UCSF1/NHS	i (0.997)	119/348	C/T	0.306	1.72	(1.19-2.48)	0.004		
		Combined		2847/8105			1.36	(1.27-1.45)	2.71E-19	0.03	(OR=1.56, $P=2.48\times10^{-10}$)[29]
rs2647012	32664458	NCI	i (1)	2142/6221	C/T	0.601	1.58	(1.46-1.71)	1.14E-31		
		UCSF2	g	210/746	C/T	0.637	1.35	(1.07-1.70)	0.01		
		SCALE	g	376/791	C/T	0.564	1.62	(1.34-1.96)	6.45E-07		
		UCSF1/NHS	i (0.992)	119/348	C/T	0.603	1.77	(1.25-2.51)	0.001		
		Combined		2847/8106			1.57	(1.47-1.68)	2.95E-40	0.34	(OR=1.56, $P=1.56\times10^{-21}$)[54]
rs10484561	32665420	NCI	i (1)	2141/6221	G/T	0.121	1.87	(1.68-2.07)	3.11E-31		
		UCSF2	i (1)	210/746	G/T	0.131	1.83	(1.35-2.47)	9.66E-05		
		SCALE	g	376/791	G/T	0.114	1.71	(1.31-2.23)	7.61E-05		
		UCSF1/NHS	i (1)	119/349	G/T	0.109	2.11	(1.32-3.38)	0.002		
		Combined		2846/8107			1.85	(1.69-2.03)	5.15E-40	0.02	(OR=1.95, $P=1.12\times10^{-29}$)[55]

rs2621416	32741868	NCI	g	2142/6221	C/T	0.276	1.42	(1.31-1.55)	3.21E-17		
		UCSF2	i (0.981)	210/745	C/T	0.272	1.3	(1.01-1.65)	0.04		
		SCALE	i (0.979)	376/790	C/T	0.281	1.36	(1.11-1.67)	0.003		
		UCSF1/NHS	i (0.950)	119/349	C/T	0.262	1.59	(1.08-2.35)	0.02		
		Combined		2847/8105			1.41	(1.31-1.51)	4.96E-21	0.60	(OR=1.57, $P=2.41\times10^{-9}$)[29]
rs241447	32796751	NCI	g	2142/6221	C/T	0.269	1.32	(1.21-1.43)	3.55E-11		
		UCSF2	g	210/746	C/T	0.265	1.34	(1.06-1.71)	0.02		
		SCALE	g	376/791	C/T	0.243	1.19	(0.96-1.48)	0.11		
		UCSF1/NHS	g	119/349	C/T	0.274	1.6	(1.10-2.32)	0.01		
		Combined		2847/8107			1.32	(1.22-1.41)	5.48E-14	0.67	(OR=1.82, $P=6.9\times10^{-8}$)[56]
DRB1*0101		NCI	i (0.975)	2142/6221	P/A	0.107	1.92	(1.71-2.16)	1.32E-28		
		UCSF2	i (1.000)	210/746	P/A	0.102	1.68	(1.20-2.36)	0.003		
		SCALE	i (1.005)	376/791	P/A	0.114	1.66	(1.26-2.20)	0.0004		
		UCSF1/NHS	i (1.009)	119/349	P/A	0.090	1.84	(1.07-3.17)	0.03		
		Combined		2847/8107	P/A		1.86	(1.69-2.06)	4.28E-34	5.45E-03	(OR=2.14, $P<0.001$)[57]
DPB1*0301		NCI	i (0.898)	2142/6221	P/A	0.102	0.49	(0.40-0.60)	6.78E-13		
		UCSF2	i (0.952)	210/746	P/A	0.094	0.55	(0.30-1.02)	0.06		
		SCALE	i (0.899)	376/791	P/A	0.094	0.59	(0.39-0.91)	0.02		
		UCSF1/NHS	i (0.785)	119/349	P/A	0.110	0.7	(0.33-1.49)	0.35		
		Combined		2847/8107	P/A		0.52	(0.44-0.61)	6.83E-15	5.58E-05	(OR=0.39, $P=4.61\times10^{-4}$)[58]
DRB1*13 *(*1301, *1302, *1303)*		NCI	i (0.985)	2142/6221	P/A	0.112	0.55	(0.48-0.63)	8.26E-18		
		UCSF2	i (0.984)	210/746	P/A	0.096	0.7	(0.46-1.05)	0.08		
		SCALE	i (0.987)	376/791	P/A	0.119	0.53	(0.38-0.74)	0.0002		
		UCSF1/NHS	i (0.906)	119/349	P/A	0.106	0.59	(0.33-1.08)	0.09		
		Combined		2847/8107	P/A		0.56	(0.50-0.63)	5.79E-22	4.91E-05	(OR=0.48, $P=0.008$)[58]
AA_DRB1_13 32660109_YF		NCI	i (0.998)	2142/6221	P/A	0.297	1.96	(1.81-2.13)	1.20E-60		

	UCSF2	i (1.024)	210/746	P/A	0.304	1.49	(1.18-1.87)	0.0007		
	SCALE	i (0.966)	376/791	P/A	0.264	1.67	(1.35-2.06)	2.26E-06		
	UCSF1/NHS	i (1.028)	119/349	P/A	0.292	2.18	(1.53-3.12)	1.89E-05		
	Combined		2847/8107	P/A		1.88	(1.76-2.02)	1.61E-70	0.37	(OR=1.76, $P=2.00\times10^{-14}$)[59]
AA_DRB1_13 32660109_SR	NCI	i (1.031)	2142/6221	P/A	0.491	0.56	(0.52-0.60)	1.37E-51		
	UCSF2	i (0.981)	210/746	P/A	0.477	0.75	(0.60-0.94)	0.01312		
	SCALE	i (1.012)	376/791	P/A	0.463	0.57	(0.47-0.69)	1.38E-08		
	UCSF1/NHS	i (1.012)	119/349	P/A	0.502	0.48	(0.34-0.68)	4.43E-05		
	Combined		2847/8107	P/A		0.57	(0.53-0.61)	5.11E-62	3.40E-01	(OR=0.60, $P=6.51\times10^{-14}$)[59]

Table S11. HLA SNPs, HLA alleles and amino acids previously reported as associated with risk of follicular lymphoma and p-values after conditioning on DRB1 position 28, rs17203612 and rs3130437.

PROXY	SNP IN LD	POS	A1/A2	r^2 SNP-PROXY	PROBE	*P*	FDR	rho
rs12194148	rs12194148	6:32444198	G/T	NA	*HLA-DRB6:NR_001298*	2.98E-15	1.37E-12	-0.9
rs12194148	rs12194148	6:32444198	G/T	NA	*HLA-DRB1:NM_002124*	9.58E-07	9.77E-05	0.68
rs12194148	rs12194148	6:32444198	G/T	NA	*HLA-DQA2:NM_020056*	2.24E-05	1.18E-03	-0.61
rs12194148	rs12194148	6:32444198	G/T	NA	*HLA-DQB1:NM_002123*	6.45E-05	3.02E-03	0.58
rs12194148	rs12194148	6:32444198	G/T	NA	*HLA-DQA1:NM_002122*	2.66E-04	1.01E-02	0.54
rs12194148	rs12194148	6:32444198	G/T	NA	*HLA-DRB5:NM_002125*	5.26E-04	1.77E-02	0.52
rs12194148	rs12194148	6:32444198	G/T	NA	*C6orf25:NM_138274*	1.05E-03	2.91E-02	0.49
rs12194148	rs12194148	6:32444198	G/T	NA	*C6orf25:NM_138275*	1.05E-03	2.91E-02	0.49
rs12194148	rs1964995	6:32449411	T/C	1	*HLA-DRB6:NR_001298*	2.48E-15	1.37E-12	-0.91
rs12194148	rs1964995	6:32449411	T/C	1	*HLA-DRB1:NM_002124*	5.60E-06	4.28E-04	0.66
rs12194148	rs1964995	6:32449411	T/C	1	*HLA-DQA2:NM_020056*	6.83E-06	4.75E-04	-0.66
rs12194148	rs1964995	6:32449411	T/C	1	*HLA-DQB1:NM_002123*	1.23E-04	5.55E-03	0.58
rs12194148	rs1964995	6:32449411	T/C	1	*HLA-DRB5:NM_002125*	5.40E-04	1.80E-02	0.53
rs12194148	rs1964995	6:32449411	T/C	1	*HLA-DQA1:NM_002122*	5.51E-04	1.81E-02	0.53
rs12194148	rs1964995	6:32449411	T/C	1	*C6orf25:NM_138274*	1.64E-03	4.36E-02	0.49
rs12194148	rs1964995	6:32449411	T/C	1	*C6orf25:NM_138275*	1.64E-03	4.36E-02	0.49
rs12194148	rs3998157	6:32678477	C/A	0.86	*HLA-DRB6:NR_001298*	6.04E-11	1.32E-08	-0.87
rs12194148	rs3998157	6:32678477	C/A	0.86	*HLA-DRB1:NM_002124*	1.23E-06	1.22E-04	0.73
rs12194148	rs3998157	6:32678477	C/A	0.86	*HLA-DQA2:NM_020056*	2.65E-05	1.36E-03	-0.66
rs12194148	**rs3998157***	6:32678477	C/A	0.86	*HLA-DQB1:NM_002123*	3.12E-05	1.59E-03	0.66
rs12194148	rs3998157	6:32678477	C/A	0.86	*HLA-DQA1:NM_002122*	7.86E-05	3.64E-03	0.63
rs12194148	rs3998157	6:32678477	C/A	0.86	*HLA-DRB5:NM_002125*	6.90E-04	2.21E-02	0.56
rs12194148	rs4273729	6:32678597	C/G	0.85	*HLA-DRB6:NR_001298*	9.79E-13	3.00E-10	-0.86
rs12194148	rs4273729	6:32678597	C/G	0.85	*HLA-DRB1:NM_002124*	4.92E-07	7.28E-05	0.69
rs12194148	**rs4273729***	6:32678597	C/G	0.85	*HLA-DQB1:NM_002123*	7.84E-07	9.77E-05	0.68
rs12194148	rs4273729	6:32678597	C/G	0.85	*HLA-DQA1:NM_002122*	7.26E-06	4.75E-04	0.64
rs12194148	rs4273729	6:32678597	C/G	0.85	*HLA-DQA2:NM_020056*	7.35E-06	4.75E-04	-0.64
rs12194148	rs4273729	6:32678597	C/G	0.85	*HLA-DRB5:NM_002125*	2.51E-04	1.01E-02	0.54
rs12194148	rs4273729	6:32678597	C/G	0.85	*HLA-DQB2:NR_003937*	3.36E-04	1.22E-02	-0.53
rs12194148	rs4410767	6:32448129	T/C	1	*HLA-DRB6:NR_001298*	2.98E-15	1.37E-12	-0.9
rs12194148	rs4410767	6:32448129	T/C	1	*HLA-DRB1:NM_002124*	9.58E-07	9.77E-05	0.68
rs12194148	rs4410767	6:32448129	T/C	1	*HLA-DQA2:NM_020056*	2.24E-05	1.18E-03	-0.61
rs12194148	rs4410767	6:32448129	T/C	1	*HLA-DQB1:NM_002123*	6.45E-05	3.02E-03	0.58
rs12194148	rs4410767	6:32448129	T/C	1	*HLA-DQA1:NM_002122*	2.66E-04	1.01E-02	0.54

rs12194148	rs4410767	6:32448129	T/C	1	*HLA-DRB5:NM_002125*	5.26E-04	1.77E-02	0.52
rs12194148	rs4410767	6:32448129	T/C	1	*C6orf25:NM_138274*	1.05E-03	2.91E-02	0.49
rs12194148	rs4410767	6:32448129	T/C	1	*C6orf25:NM_138275*	1.05E-03	2.91E-02	0.49
rs12194148	rs5007260	6:32379047	G/A	0.8	*HLA-DRB6:NR_001298*	5.40E-12	1.38E-09	-0.85
rs12194148	rs5007260	6:32379047	G/A	0.8	*HLA-DRB1:NM_002124*	1.25E-06	1.22E-04	0.68
rs12194148	rs5007260	6:32379047	G/A	0.8	*HLA-DQA2:NM_020056*	1.23E-05	7.60E-04	-0.63
rs12194148	rs5007260	6:32379047	G/A	0.8	*HLA-DQB1:NM_002123*	1.01E-04	4.62E-03	0.58
rs12194148	rs5007260	6:32379047	G/A	0.8	*HLA-DQA1:NM_002122*	1.34E-04	5.96E-03	0.57
rs12194148	rs5007260	6:32379047	G/A	0.8	*HLA-DRB5:NM_002125*	2.75E-04	1.03E-02	0.55
rs12194148	rs5020946	6:32450089	G/T	1	*HLA-DRB6:NR_001298*	2.98E-15	1.37E-12	-0.9
rs12194148	rs5020946	6:32450089	G/T	1	*HLA-DRB1:NM_002124*	9.58E-07	9.77E-05	0.68
rs12194148	rs5020946	6:32450089	G/T	1	*HLA-DQA2:NM_020056*	2.24E-05	1.18E-03	-0.61
rs12194148	rs5020946	6:32450089	G/T	1	*HLA-DQB1:NM_002123*	6.45E-05	3.02E-03	0.58
rs12194148	rs5020946	6:32450089	G/T	1	*HLA-DQA1:NM_002122*	2.66E-04	1.01E-02	0.54
rs12194148	rs5020946	6:32450089	G/T	1	*HLA-DRB5:NM_002125*	5.26E-04	1.77E-02	0.52
rs12194148	rs5020946	6:32450089	G/T	1	*C6orf25:NM_138274*	1.05E-03	2.91E-02	0.49
rs12194148	rs5020946	6:32450089	G/T	1	*C6orf25:NM_138275*	1.05E-03	2.91E-02	0.49
rs12194148	rs6932517	6:32678182	C/G	0.84	*HLA-DRB6:NR_001298*	2.66E-11	6.10E-09	-0.84
rs12194148	rs6932517	6:32678182	C/G	0.84	*HLA-DQA2:NM_020056*	8.31E-07	9.77E-05	-0.7
rs12194148	rs6932517	6:32678182	C/G	0.84	*HLA-DRB1:NM_002124*	9.11E-07	9.77E-05	0.7
rs12194148	**rs6932517***	6:32678182	C/G	0.84	*HLA-DQB1:NM_002123*	2.78E-06	2.45E-04	0.67
rs12194148	rs6932517	6:32678182	C/G	0.84	*HLA-DQA1:NM_002122*	2.06E-05	1.18E-03	0.63
rs12194148	rs6932517	6:32678182	C/G	0.84	*HLA-DRB5:NM_002125*	4.92E-04	1.76E-02	0.53
rs12194148	rs6932517	6:32678182	C/G	0.84	*HLA-DQB2:NR_003937*	7.68E-04	2.45E-02	-0.52
rs12194148	rs6932517	6:32678182	C/G	0.84	*BTNL2:NM_019602*	9.28E-04	2.88E-02	0.51
rs12194148	rs7748270	6:32448599	C/T	0.97	*HLA-DRB6:NR_001298*	2.98E-15	1.37E-12	-0.9
rs12194148	rs7748270	6:32448599	C/T	0.97	*HLA-DRB1:NM_002124*	9.58E-07	9.77E-05	0.68
rs12194148	rs7748270	6:32448599	C/T	0.97	*HLA-DQA2:NM_020056*	2.24E-05	1.18E-03	-0.61
rs12194148	rs7748270	6:32448599	C/T	0.97	*HLA-DQB1:NM_002123*	6.45E-05	3.02E-03	0.58
rs12194148	rs7748270	6:32448599	C/T	0.97	*HLA-DQA1:NM_002122*	2.66E-04	1.01E-02	0.54
rs12194148	rs7748270	6:32448599	C/T	0.97	*HLA-DRB5:NM_002125*	5.26E-04	1.77E-02	0.52
rs12194148	rs7748270	6:32448599	C/T	0.97	*C6orf25:NM_138274*	1.05E-03	2.91E-02	0.49
rs12194148	rs7748270	6:32448599	C/T	0.97	*C6orf25:NM_138275*	1.05E-03	2.91E-02	0.49
rs12194148	rs9271586	6:32590899	G/T	0.84	*HLA-DRB6:NR_001298*	1.64E-08	3.43E-06	-0.82
rs12194148	rs9271586	6:32590899	G/T	0.84	*HLA-DRB1:NM_002124*	2.18E-04	9.25E-03	0.62
rs12194148	rs9271586	6:32590899	G/T	0.84	*HLA-DQA2:NM_020056*	1.51E-03	4.11E-02	-0.55

rs12194148	rs9271588	6:32590953	T/C	0.91	*HLA-DRB6:NR_001298*	5.28E-13	2.20E-10	-0.88
rs12194148	rs9271588	6:32590953	T/C	0.91	*HLA-DQA2:NM_020056*	1.15E-05	7.30E-04	-0.65
rs12194148	rs9271588	6:32590953	T/C	0.91	*HLA-DRB1:NM_002124*	1.21E-05	7.60E-04	0.65
rs12194148	**rs9271588***	6:32590953	T/C	0.91	*HLA-DQB1:NM_002123*	1.62E-04	7.08E-03	0.57
rs12194148	rs9271588	6:32590953	T/C	0.91	*HLA-DRB5:NM_002125*	8.46E-04	2.68E-02	0.52
rs12194148	rs9271588	6:32590953	T/C	0.91	*HLA-DQA1:NM_002122*	8.96E-04	2.80E-02	0.52
rs12194148	rs9275517	6:32674649	A/G	0.88	*HLA-DRB6:NR_001298*	8.60E-12	2.08E-09	-0.88
rs12194148	rs9275517	6:32674649	A/G	0.88	*HLA-DQA2:NM_020056*	3.85E-06	3.10E-04	-0.7
rs12194148	rs9275517	6:32674649	A/G	0.88	*HLA-DRB1:NM_002124*	3.96E-06	3.13E-04	0.7
rs12194148	**rs9275517***	6:32674649	A/G	0.88	*HLA-DQB1:NM_002123*	6.49E-06	4.75E-04	0.69
rs12194148	rs9275517	6:32674649	A/G	0.88	*HLA-DQA1:NM_002122*	2.63E-05	1.36E-03	0.65
rs12194148	rs9275517	6:32674649	A/G	0.88	*HLA-DRB5:NM_002125*	2.35E-04	9.89E-03	0.59
rs12194148	rs9275517	6:32674649	A/G	0.88	*HLA-DQB2:NR_003937*	1.03E-03	2.91E-02	-0.54
rs12194148	rs9275517	6:32674649	A/G	0.88	*BTNL2:NM_019602*	1.11E-03	3.07E-02	0.54
rs12194148	rs9275524	6:32675109	T/C	0.88	*HLA-DRB6:NR_001298*	3.63E-12	9.79E-10	-0.86
rs12194148	**rs9275524***	6:32675109	T/C	0.88	*HLA-DQB1:NM_002123*	2.12E-06	1.91E-04	0.68
rs12194148	rs9275524	6:32675109	T/C	0.88	*HLA-DRB1:NM_002124*	3.23E-06	2.75E-04	0.68
rs12194148	rs9275524	6:32675109	T/C	0.88	*HLA-DQA2:NM_020056*	3.63E-06	2.97E-04	-0.67
rs12194148	rs9275524	6:32675109	T/C	0.88	*HLA-DQA1:NM_002122*	2.16E-05	1.18E-03	0.63
rs12194148	rs9275524	6:32675109	T/C	0.88	*HLA-DRB5:NM_002125*	2.08E-04	8.92E-03	0.57
rs12194148	rs9275524	6:32675109	T/C	0.88	*HLA-DQB2:NR_003937*	6.83E-04	2.21E-02	-0.53
rs12194148	rs9275565	6:32677938	T/C	0.85	*HLA-DRB6:NR_001298*	9.79E-13	3.00E-10	-0.86
rs12194148	rs9275565	6:32677938	T/C	0.85	*HLA-DRB1:NM_002124*	4.92E-07	7.28E-05	0.69
rs12194148	**rs9275565***	6:32677938	T/C	0.85	*HLA-DQB1:NM_002123*	7.84E-07	9.77E-05	0.68
rs12194148	rs9275565	6:32677938	T/C	0.85	*HLA-DQA1:NM_002122*	7.26E-06	4.75E-04	0.64
rs12194148	rs9275565	6:32677938	T/C	0.85	*HLA-DQA2:NM_020056*	7.35E-06	4.75E-04	-0.64
rs12194148	rs9275565	6:32677938	T/C	0.85	*HLA-DRB5:NM_002125*	2.51E-04	1.01E-02	0.54
rs12194148	rs9275565	6:32677938	T/C	0.85	*HLA-DQB2:NR_003937*	3.36E-04	1.22E-02	-0.53
rs12194148	rs9275572	6:32678999	A/G	0.88	*HLA-DRB6:NR_001298*	3.63E-12	9.79E-10	-0.86
rs12194148	**rs9275572***	6:32678999	A/G	0.88	*HLA-DQB1:NM_002123*	2.12E-06	1.91E-04	0.68
rs12194148	rs9275572	6:32678999	A/G	0.88	*HLA-DRB1:NM_002124*	3.23E-06	2.75E-04	0.68
rs12194148	rs9275572	6:32678999	A/G	0.88	*HLA-DQA2:NM_020056*	3.63E-06	2.97E-04	-0.67
rs12194148	rs9275572	6:32678999	A/G	0.88	*HLA-DQA1:NM_002122*	2.16E-05	1.18E-03	0.63
rs12194148	rs9275572	6:32678999	A/G	0.88	*HLA-DRB5:NM_002125*	2.08E-04	8.92E-03	0.57
rs12194148	rs9275572	6:32678999	A/G	0.88	*HLA-DQB2:NR_003937*	6.83E-04	2.21E-02	-0.53
rs12194148	rs9275573	6:32679146	C/G	0.85	*HLA-DRB6:NR_001298*	9.79E-13	3.00E-10	-0.86

rs12194148	rs9275573	6:32679146	C/G	0.85	*HLA-DRB1:NM_002124*	4.92E-07	7.28E-05	0.69
rs12194148	**rs9275573***	6:32679146	C/G	0.85	*HLA-DQB1:NM_002123*	7.84E-07	9.77E-05	0.68
rs12194148	rs9275573	6:32679146	C/G	0.85	*HLA-DQA1:NM_002122*	7.26E-06	4.75E-04	0.64
rs12194148	rs9275573	6:32679146	C/G	0.85	*HLA-DQA2:NM_020056*	7.35E-06	4.75E-04	-0.64
rs12194148	rs9275573	6:32679146	C/G	0.85	*HLA-DRB5:NM_002125*	2.51E-04	1.01E-02	0.54
rs12194148	rs9275573	6:32679146	C/G	0.85	*HLA-DQB2:NR_003937*	3.36E-04	1.22E-02	-0.53
rs12194148	rs9378212	6:32445691	C/T	1	*HLA-DRB6:NR_001298*	2.98E-15	1.37E-12	-0.9
rs12194148	rs9378212	6:32445691	C/T	1	*HLA-DRB1:NM_002124*	9.58E-07	9.77E-05	0.68
rs12194148	rs9378212	6:32445691	C/T	1	*HLA-DQA2:NM_020056*	2.24E-05	1.18E-03	-0.61
rs12194148	rs9378212	6:32445691	C/T	1	*HLA-DQB1:NM_002123*	6.45E-05	3.02E-03	0.58
rs12194148	rs9378212	6:32445691	C/T	1	*HLA-DQA1:NM_002122*	2.66E-04	1.01E-02	0.54
rs12194148	rs9378212	6:32445691	C/T	1	*HLA-DRB5:NM_002125*	5.26E-04	1.77E-02	0.52
rs12194148	rs9378212	6:32445691	C/T	1	*C6orf25:NM_138274*	1.05E-03	2.91E-02	0.49
rs12194148	rs9378212	6:32445691	C/T	1	*C6orf25:NM_138275*	1.05E-03	2.91E-02	0.49
rs12194148	rs9378213	6:32448398	T/G	1	*HLA-DRB6:NR_001298*	2.98E-15	1.37E-12	-0.9
rs12194148	rs9378213	6:32448398	T/G	1	*HLA-DRB1:NM_002124*	9.58E-07	9.77E-05	0.68
rs12194148	rs9378213	6:32448398	T/G	1	*HLA-DQA2:NM_020056*	2.24E-05	1.18E-03	-0.61
rs12194148	rs9378213	6:32448398	T/G	1	*HLA-DQB1:NM_002123*	6.45E-05	3.02E-03	0.58
rs12194148	rs9378213	6:32448398	T/G	1	*HLA-DQA1:NM_002122*	2.66E-04	1.01E-02	0.54
rs12194148	rs9378213	6:32448398	T/G	1	*HLA-DRB5:NM_002125*	5.26E-04	1.77E-02	0.52
rs12194148	rs9378213	6:32448398	T/G	1	*C6orf25:NM_138274*	1.05E-03	2.91E-02	0.49
rs12194148	rs9378213	6:32448398	T/G	1	*C6orf25:NM_138275*	1.05E-03	2.91E-02	0.49
rs12194148	rs9378264	6:32443451	G/A	1	*HLA-DRB6:NR_001298*	2.48E-15	1.37E-12	-0.91
rs12194148	rs9378264	6:32443451	G/A	1	*HLA-DRB1:NM_002124*	5.60E-06	4.28E-04	0.66
rs12194148	rs9378264	6:32443451	G/A	1	*HLA-DQA2:NM_020056*	6.83E-06	4.75E-04	-0.66
rs12194148	rs9378264	6:32443451	G/A	1	*HLA-DQB1:NM_002123*	1.23E-04	5.55E-03	0.58
rs12194148	rs9378264	6:32443451	G/A	1	*HLA-DRB5:NM_002125*	5.40E-04	1.80E-02	0.53
rs12194148	rs9378264	6:32443451	G/A	1	*HLA-DQA1:NM_002122*	5.51E-04	1.81E-02	0.53
rs12194148	rs9378264	6:32443451	G/A	1	*C6orf25:NM_138274*	1.64E-03	4.36E-02	0.49
rs12194148	rs9378264	6:32443451	G/A	1	*C6orf25:NM_138275*	1.64E-03	4.36E-02	0.49
rs12194148	rs9378266	6:32448189	G/T	1	*HLA-DRB6:NR_001298*	2.98E-15	1.37E-12	-0.9
rs12194148	rs9378266	6:32448189	G/T	1	*HLA-DRB1:NM_002124*	9.58E-07	9.77E-05	0.68
rs12194148	rs9378266	6:32448189	G/T	1	*HLA-DQA2:NM_020056*	2.24E-05	1.18E-03	-0.61
rs12194148	rs9378266	6:32448189	G/T	1	*HLA-DQB1:NM_002123*	6.45E-05	3.02E-03	0.58
rs12194148	rs9378266	6:32448189	G/T	1	*HLA-DQA1:NM_002122*	2.66E-04	1.01E-02	0.54
rs12194148	rs9378266	6:32448189	G/T	1	*HLA-DRB5:NM_002125*	5.26E-04	1.77E-02	0.52

rs12194148	rs9378266	6:32448189	G/T	1	*C6orf25:NM_138274*	1.05E-03	2.91E-02	0.49
rs12194148	rs9378266	6:32448189	G/T	1	*C6orf25:NM_138275*	1.05E-03	2.91E-02	0.49
rs12194148	rs9380318	6:32677669	C/T	0.85	*HLA-DRB6:NR_001298*	9.79E-13	3.00E-10	-0.86
rs12194148	rs9380318	6:32677669	C/T	0.85	*HLA-DRB1:NM_002124*	4.92E-07	7.28E-05	0.69
rs12194148	**rs9380318***	6:32677669	C/T	0.85	*HLA-DQB1:NM_002123*	7.84E-07	9.77E-05	0.68
rs12194148	rs9380318	6:32677669	C/T	0.85	*HLA-DQA1:NM_002122*	7.26E-06	4.75E-04	0.64
rs12194148	rs9380318	6:32677669	C/T	0.85	*HLA-DQA2:NM_020056*	7.35E-06	4.75E-04	-0.64
rs12194148	rs9380318	6:32677669	C/T	0.85	*HLA-DRB5:NM_002125*	2.51E-04	1.01E-02	0.54
rs12194148	rs9380318	6:32677669	C/T	0.85	*HLA-DQB2:NR_003937*	3.36E-04	1.22E-02	-0.53
rs12194148	rs9391786	6:32448561	A/G	1	*HLA-DRB6:NR_001298*	2.98E-15	1.37E-12	-0.9
rs12194148	rs9391786	6:32448561	A/G	1	*HLA-DRB1:NM_002124*	9.58E-07	9.77E-05	0.68
rs12194148	rs9391786	6:32448561	A/G	1	*HLA-DQA2:NM_020056*	2.24E-05	1.18E-03	-0.61
rs12194148	rs9391786	6:32448561	A/G	1	*HLA-DQB1:NM_002123*	6.45E-05	3.02E-03	0.58
rs12194148	rs9391786	6:32448561	A/G	1	*HLA-DQA1:NM_002122*	2.66E-04	1.01E-02	0.54
rs12194148	rs9391786	6:32448561	A/G	1	*HLA-DRB5:NM_002125*	5.26E-04	1.77E-02	0.52
rs12194148	rs9391786	6:32448561	A/G	1	*C6orf25:NM_138274*	1.05E-03	2.91E-02	0.49
rs12194148	rs9391786	6:32448561	A/G	1	*C6orf25:NM_138275*	1.05E-03	2.91E-02	0.49
rs3130439	rs11962994	6:31219289	A/G	0.94	*PSORS1C3:NR_026816*	3.18E-07	5.40E-05	0.73
rs3130439	rs11967600	6:31199573	T/C	0.82	*PSORS1C3:NR_026816*	1.60E-04	7.05E-03	0.61
rs3130439	rs11967600	6:31199573	T/C	0.82	*PSORS1C2:NM_014069*	3.52E-04	1.27E-02	0.58
rs3130439	rs2394892	6:31205382	A/G	0.89	*PSORS1C3:NR_026816*	2.79E-07	4.93E-05	0.72
rs3130439	rs28397299	6:31207704	A/G	0.83	*PSORS1C3:NR_026816*	1.64E-05	9.89E-04	0.64
rs3130439	rs3130439	6:31221023	G/A	NA	*PSORS1C3:NR_026816*	1.32E-05	8.10E-04	0.65
rs3130439	rs3130439	6:31221023	G/A	NA	*PSORS1C2:NM_014069*	1.39E-03	3.81E-02	0.51
rs3130439	rs3130439	6:31221023	G/A	NA	*DPCR1:NM_080870*	1.55E-03	4.21E-02	-0.5
rs3130439	rs3869117	6:31205923	G/C	0.93	*PSORS1C3:NR_026816*	1.38E-06	1.31E-04	0.68
rs3130439	rs3869117	6:31205923	G/C	0.93	*PSORS1C2:NM_014069*	8.94E-04	2.80E-02	0.5
rs3130439	rs6457350	6:31204109	C/T	0.87	*PSORS1C3:NR_026816*	1.79E-06	1.68E-04	0.68
rs3130439	rs6908994	6:31198709	T/C	0.89	*PSORS1C3:NR_026816*	2.79E-07	4.93E-05	0.72
rs3130439	rs7745906	6:31204008	A/G	0.89	*PSORS1C3:NR_026816*	2.77E-07	4.93E-05	0.73
rs3130439	rs7768431	6:31202665	A/G	0.89	*PSORS1C3:NR_026816*	2.36E-07	4.71E-05	0.71

A1/A2 = minor/major allele in HapMap-CEU r28. Gene expression changes were estimated for the minor allele.

*The highlighted eQTLs also showed significant correlation with lower methylation levels in the same gene.

Table S12. Results from the eQTL analysis of rs12194148 and rs3130439, proxies, respectively, for the independent markers in the HLA forward stepwise analysis, rs17203612 and rs3130437. eQTL analysis was performed using a publicly available RNA Seq dataset (GEO accession number GSE16921) containing

whole-genome gene expression data in transformed lymphoblastoid cell lines from 41 HapMap-CEU samples. Whole-genome genotyping data for the same HapMap-CEU individuals (release #28) were directly downloaded from HapMap. As the two SNPs that remained statistically significant in the stepwise conditional analysis, rs17203612 and rs3130439, were not available in HapMap, rs12194148 (r^2 = 0.98 with rs17203612) and rs3130439 (r^2= 0.93 with rs3130437) were selected as proxies, and the eQTL analysis was conducted on these two proxies and SNPs in LD (r^2> 0.8) by correlating genotype and expression levels of probes within 1Mb of the SNPs. Correlation between expression and genotype for each SNP-probe pair was tested using the Spearman's rank test with t-distribution approximation and were estimated with respect to the minor allele in HapMap-CEU. *P*-values were adjusted for multiple comparisons using the Benjamini-Hochberg false-discovery rate (FDR) and eQTL were considered statistically significant at an FDR *P*-value threshold < 0.05.

PROXY	SNP in LD	POS	A1/A2[a]	PROBE	GENE	RHO	*P*	FDR
rs12194148	**rs3998157***	6:32678477	C/A	cg01889448	HLA-DQB1	-0.48	7.15E-06	3.86E-05
rs12194148	**rs4273729***	6:32678597	C/G	cg01889448	HLA-DQB1	-0.45	8.53E-06	4.19E-05
rs12194148	**rs6932517***	6:32678182	C/G	cg01889448	HLA-DQB1	-0.46	1.36E-05	5.46E-05
rs12194148	rs9271586	6:32590899	G/T	cg01889448	HLA-DQB1	-0.48	1.42E-05	5.46E-05
rs12194148	**rs9271588***	6:32590953	T/C	cg01889448	HLA-DQB1	-0.47	9.33E-06	4.20E-05
rs12194148	**rs9275517***	6:32674649	A/G	cg01889448	HLA-DQB1	-0.47	1.63E-05	5.86E-05
rs12194148	**rs9275524***	6:32675109	T/C	cg01889448	HLA-DQB1	-0.45	2.43E-05	7.72E-05
rs12194148	**rs9275565***	6:32677938	T/C	cg01889448	HLA-DQB1	-0.47	3.74E-06	2.79E-05
rs12194148	**rs9275572***	6:32678999	A/G	cg01889448	HLA-DQB1	-0.45	2.43E-05	7.72E-05
rs12194148	**rs9275573***	6:32679146	C/G	cg01889448	HLA-DQB1	-0.47	3.74E-06	2.79E-05
rs12194148	**rs9380318***	6:32677669	C/T	cg01889448	HLA-DQB1	-0.47	4.12E-06	2.79E-05
rs3130439	rs11962994	6:31219289	A/G	cg10409680	HLA-C	0.43	6.10E-05	1.65E-04
rs3130439	rs11967600	6:31199573	T/C	cg10409680	HLA-C	0.4	4.70E-04	1.21E-03
rs3130439	rs2394892	6:31205382	A/G	cg10409680	HLA-C	0.49	3.98E-06	2.79E-05
rs3130439	rs28397299	6:31207704	A/G	cg10409680	HLA-C	0.49	4.13E-06	2.79E-05
rs3130439	rs3130439	6:31221023	G/A	cg10409680	HLA-C	0.43	3.28E-05	9.85E-05
rs3130439	rs3869117	6:31205923	G/C	cg10409680	HLA-C	0.47	2.58E-06	2.79E-05
rs3130439	rs6457350	6:31204109	C/T	cg10409680	HLA-C	0.43	3.56E-05	1.01E-04
rs3130439	rs6908994	6:31198709	T/C	cg10409680	HLA-C	0.49	3.98E-06	2.79E-05
rs3130439	rs7745906	6:31204008	A/G	cg10409680	HLA-C	0.5	3.25E-06	2.79E-05
rs3130439	rs7768431	6:31202665	A/G	cg10409680	HLA-C	0.46	5.37E-06	3.22E-05

[a]A1/A2 = minor/major allele in HapMap-CEU r28. Methylation changes were estimated for the minor allele.

*The highlighted meQTLs also showed significant correlation with higher expression levels in the same gene

Table S13. Results from the meQTL analysis of rs12194148 and rs3130439, proxies, respectively, for the independent markers in the conditional analysis, rs17203612 and rs3130437. meQTL analysis was performed using a publicly available dataset (GEO accession number GSE27146) that contained 27,578 DNA methylation measurements near the transcription start sites of 14,000 genes in 180 HapMap samples. Methylation probes were mapped to the human genome sequence (hg19) using BLAT and those that mapped to multiple locations with up to two mismatches were discarded, leaving 26,375 probes for analysis. Only the 90 samples of CEU origin were used in this study, and only SNPs and methylation probes that were located within 50kb were tested for association. Correlation between SNP genotypes and methylation levels was tested using the Spearman's rank correlation test and estimated with respect to the minor allele in HapMap-CEU. meQTL were considered significant at an FDR adjusted *P*-value < 0.05.

SNP	Gene transcript	Effect allele	Other allele	Beta for NHL SNP[a]	*P* for NHL SNP[a]	*P* for NHL SNP conditioned on peak SNP[b]	Peak SNP for transcript[c]	Beta for Peak SNP[d]	*P* for Peak SNP[d]	*P* for peak SNP conditioned on NHL SNP[e]
rs3130437	*HCG22*	C	A	-2.215	2.03E-12	0.82	rs116195588	-2.034	7.56E-27	4.58E-20
rs3130437	*HCG27*	C	A	1.650	7.10E-10	0.42	rs116794933	-0.615	5.86E-14	2.36E-07
rs3130437	*HLA-C*	C	A	0.882	1.13E-07	0.007	rs116398710	0.556	1.21E-11	1.13E-07
rs3130437	*HLA-C*	C	A	0.760	4.24E-07	0.08	rs149683222	-0.198	3.85E-10	3.11E-05
rs3130437	*TCF19*	C	A	0.847	5.74E-06	0.002	rs140242258	0.830	1.05E-28	8.51E-28
rs3130437	*HLA-B*	C	A	0.617	2.53E-05	0.002	rs140242258	0.461	3.51E-15	1.03E-13

[a]Beta and p-value for the association between the NHL SNP and gene transcript.
[b]p-value for the association between the NHL SNP and gene transcript after adjustment for the peak SNP
[c]Peak SNP is the most significant SNP associated with the gene transcript
[d]Beta and p-value for the association between the peak SNP and the gene transcript
[e]P-value for the association between the peak SNP and the gene transcript after adjustment for the NHL SNP

Table S14. Expression quantitative trait loci (eQTL) associations with FDR < 1% from the childhood asthma dataset in the HLA region. eQTL analysis was conducted using a publicly available childhood asthma microarray dataset (GEO accession number GSE8052). As described previously for this dataset, peripheral blood lymphocytes were transformed into lymphoblastoid cell lines for 830 parents and offspring from 206 families of European ancestry. Using extracted RNA, gene expression was assessed with the Affymetrix HG-U133 Plus 2.0 chip. Genotyping was conducted using the Illumina Human1M Beadchip and Illumina HumanHap300K Beadchip, and imputation performed using data from 1000 Genomes Project. All SNPs selected for replication were tested for *cis* associations (defined as gene transcripts within 1 Mb), assuming an additive genetic model, adjusting for non-genetic effects in the gene expression value. To gain insight into the relative importance of eQTL associations with our SNPs compared to other SNPs in the region with stronger eQTL associations, we also conducted conditional analyses, in which both the FL SNP and the most significant SNP for the particular gene transcript (i.e., the peak SNP) were included in the same model. Only *cis* associations that reached $P<6.8\times10^{-5}$, which corresponds to a FDR of 1% are reported.

chr	pos (hg19)	LD (r^2)	LD (D')	variant	Ref	Alt	EUR freq	Enhancer Histone marks	DNAse	Proteins bound	eQTL tissues	Motifs changed	GENCODE genes	dbSNP func annot
6	31220895	1	1	rs3130437	A	C	0.61					4 altered motifs	16kb 3' of HLA-C	
6	31221153	0.93	0.98	rs2394946	G	A	0.61					Ncx,Pou1f1,Sox	15kb 3' of HLA-C	
6	32442836	0.83	0.97	rs13211921	T	G	0.43					13 altered motifs	30kb 3' of HLA-DRA	
6	32443172	0.92	0.97	rs9391879	C	T	0.45					SZF1-1,Spz1	30kb 3' of HLA-DRA	
6	32443451	0.89	0.95	rs9378264	G	A	0.45	K562					31kb 3' of HLA-DRA	
6	32443666	0.8	0.97	rs28895242	G	C	0.42					CEBPA,STAT	31kb 3' of HLA-DRA	
6	32444198	0.92	0.97	rs12194148	G	T	0.45						31kb 3' of HLA-DRA	
6	32444330	0.9	0.97	rs12207473	A	C	0.44					7 altered motifs	32kb 3' of HLA-DRA	
6	32444544	0.89	0.94	rs12195582	C	T	0.46						32kb 3' of HLA-DRA	
6	32444733	0.9	0.96	rs113660101	TG	T	0.45					GATA,SIX5	32kb 3' of HLA-DRA	
6	32445079	0.92	0.98	rs28895255	C	T	0.45					NF-kappaB,Rad21	32kb 3' of HLA-DRA	
6	32445114	0.87	0.98	rs28895257	A	G	0.43					Foxp3	32kb 3' of HLA-DRA	
6	32445117	0.87	0.98	rs28895258	C	T	0.43					AP-1,p53	32kb 3' of HLA-DRA	
6	32445306	0.88	0.99	rs28895261	G	A	0.43					AhR::Arnt,Arnt,BDP1	32kb 3' of HLA-DRA	
6	32445691	0.92	0.96	rs9378212	C	T	0.46		MCF-7	CTCF		5 altered motifs	33kb 3' of HLA-DRA	
6	32445992	0.91	0.98	rs17209866	G	A	0.44					Pitx2	33kb 3' of HLA-DRA	
6	32446010	0.91	0.96	rs17209873	T	C	0.46					Pbx-1,Rhox11	33kb 3' of HLA-DRA	
6	32446051	0.93	0.98	rs17209887	A	C	0.45						33kb 3' of HLA-DRA	
6	32446071	0.95	0.98	rs17203514	C	G	0.45					Hoxa5	33kb 3' of HLA-DRA	
6	32446307	0.86	0.96	rs17203549	T	C	0.44					Pax-5	33kb 3' of HLA-DRA	
6	32446425	0.83	0.98	rs139916710	GAAT	G	0.42					4 altered motifs	34kb 3' of HLA-DRA	

6	32446459	0.86	0.99	rs17203563	G	A, C, T	0.42				34kb 3' of HLA-DRA
6	32446853	1	1	rs17203612	A	T	0.46			8 altered motifs	34kb 3' of HLA-DRA
6	32446922	0.89	1	rs17203619	T	C	0.43				34kb 3' of HLA-DRA
6	32446994	0.81	1	rs17203626	G	A, T	0.41				34kb 3' of HLA-DRA
6	32447014	0.95	1	rs17203636	G	A	0.45			Pou3f2,Sox	34kb 3' of HLA-DRA
6	32447111	0.95	1	rs29001652	A	G	0.45		H7-hESC	Glis2,NERF1a	34kb 3' of HLA-DRA
6	32447162	0.96	0.99	rs29001478	G	A	0.45		H7-hESC	4 altered motifs	34kb 3' of HLA-DRA
6	32447216	0.96	0.99	rs29001620	A	T	0.45		H7-hESC	NRSF,Sin3Ak-20	34kb 3' of HLA-DRA
6	32447219	0.96	0.99	rs28732246	G	A	0.45		H7-hESC	NRSF,PU.1,Sin3Ak-20	34kb 3' of HLA-DRA
6	32447341	0.96	0.99	rs29001568	G	A	0.45			Maf,Nkx2,RXRA	35kb 3' of HLA-DRA
6	32447715	0.92	0.99	rs4994859	A	G	0.44			4 altered motifs	35kb 3' of HLA-DRA
6	32447873	0.92	0.99	rs4994855	A	G	0.44			6 altered motifs	35kb 3' of HLA-DRA
6	32447900	0.91	0.99	rs4994854	T	C	0.44			Pax-4,Sox	35kb 3' of HLA-DRA
6	32448129	0.81	0.99	rs4410767	T	C	0.41			GATA,Irf	35kb 3' of HLA-DRA
6	32449138	0.9	0.97	rs9394098	C	T	0.44			Smad3	36kb 3' of HLA-DRB5
6	32449160	0.87	0.97	rs9394099	G	T	0.44			6 altered motifs	36kb 3' of HLA-DRB5
6	32449188	0.86	0.96	rs9394100	G	A	0.47			Pou5f1	36kb 3' of HLA-DRB5
6	32449198	0.85	0.94	rs9394101	A	G	0.46			Irf,TEF	36kb 3' of HLA-DRB5
6	32449293	0.87	0.97	rs5018660	A	G	0.44			Dmbx1,Mef2,OTX	36kb 3' of HLA-DRB5
6	32449411	0.85	0.97	rs1964995	T	C	0.43			5 altered motifs	36kb 3' of HLA-DRB5
6	32450089	0.87	0.98	rs5020946	G	T	0.43	GM12878	Monocytes-CD14+_RO01746	AP-2,STAT	35kb 3' of HLA-DRB5
6	32450678	0.83	0.96	rs34452456	T	C	0.44	GM12878		Pax-4	34kb 3' of HLA-DRB5
6	32451570	0.81	0.92	rs13194770	C	T	0.45			5 altered motifs	34kb 3' of HLA-DRB5

Table S15. Results from HaploReg analysis of HLA follicular lymphoma risk loci and their correlated (r^2 > 0.80) surrogates.

3. **SUPPLEMENTAL NOTES**

NOTE S1: ACKNOWLEDGEMENTS

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).

BC (J.S., A.B.W.) – Canadian Institutes for Health Research (CIHR). Canadian Cancer Society. Michael Smith Foundation for Health Research.

CPSII (L.T.) – 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 (E.R.) - Leukaemia & Lymphoma Research.

ENGELA (J.C.) – 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 (E.R.) – 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.

HPFS (Walter C. Willet) – The HPFS was supported in part by National Institutes of Health grants CA167552, CA149445, CA098122, CA098566 (K.A.B.), and K07 CA115687 (B.M.B.). 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. In addition, this study was approved by the Connecticut Department of Public Health (DPH) Human Investigations Committee. Certain data used in this publication were obtained from the DPH. The authors assume full responsibility for analyses and interpretation of these data.

Iowa-Mayo SPORE (G.W., J.R.C., T.E.W.) – National Institutes of Health (CA97274). Specialized Programs of Research Excellence (SPORE) in Human Cancer (P50 CA97274). Molecular Epidemiology of Non-Hodgkin Lymphoma Survival (R01 CA129539).

Italian GxE (P.C.) - 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.

LYSA (G.S, H.G) – Institut National du Cancer (INCa, Paris) grant 2008-020

Mayo Clinic Case-Control (J.R.C.) – National Institutes of Health (R01 CA92153).

MCCS (G.G.G., G.S.) – 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 (X.W.) - Institutional support to the Center for Translational and Public Health Genomics.

MSKCC (K.O.) – 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.

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 (Meir J. Stampfer) – The NHS was supported in part by National Institutes of Health grants CA87969, CA49449, CA149445, CA098122, CA134958, CA098566 (K.A.B.), and K07 CA115687 (B.M.B.). 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. In addition, this study was approved by the Connecticut Department of Public Health (DPH) Human Investigations Committee. Certain data used in this publication were obtained from the DPH. The authors assume full responsibility for analyses and interpretation of these data.

NSW (C.M.Vajdic) - 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 (K.E.S., H.O.A., H.H.) – 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/UCSF2 (C.F.S.)- These studies were supported by the NCI, National Institutes of Health, grants CA1046282 and CA154643 (C.F.S.), and R01CA87014, R01CA45614, R03CA14397, and R03CA150037 (E.A.H., P.M.B). 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 – WHI investigators are: *Program Office* - (National Heart, Lung, and Blood Institute, Bethesda, Maryland) Jacques Rossouw, Shari Ludlam, Dale Burwen, Joan McGowan, Leslie Ford, and Nancy Geller; *Clinical Coordinating Center* - (Fred Hutchinson Cancer Research Center, Seattle, WA) Garnet Anderson, Ross Prentice, Andrea LaCroix, and Charles Kooperberg; *Investigators and Academic Centers* - (Brigham and Women's Hospital, Harvard Medical School, Boston, MA) JoAnn E. Manson; (MedStar Health Research Institute/Howard University, Washington, DC) Barbara V. Howard; (Stanford Prevention Research Center, Stanford, CA) Marcia L. Stefanick; (The Ohio State University, Columbus, OH) Rebecca Jackson; (University of Arizona, Tucson/Phoenix, AZ) Cynthia A. Thomson; (University at Buffalo, Buffalo, NY) Jean Wactawski-Wende; (University of Florida, Gainesville/Jacksonville, FL) Marian Limacher; (University of Iowa, Iowa City/Davenport, IA) Robert Wallace; (University of Pittsburgh, Pittsburgh, PA) Lewis Kuller; (Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker; *Women's Health Initiative Memory Study* - (Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker. 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 HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C.

YALE (T.Z.) – National Cancer Institute (CA62006).

4. SUPPLEMENTAL REFERENCES

1. (1994). The alpha-tocopherol, beta-carotene lung cancer prevention study: design, methods, participant characteristics, and compliance. The ATBC Cancer Prevention Study Group. Annals of epidemiology 4, 1-10.
2. Calle, E.E., Rodriguez, C., Jacobs, E.J., Almon, M.L., Chao, A., McCullough, M.L., Feigelson, H.S., and Thun, M.J. (2002). The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. Cancer 94, 500-511.
3. Riboli, E., and Kaaks, R. (1997). The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. International journal of epidemiology 26 Suppl 1, S6-14.
4. Riboli, E., Hunt, K.J., Slimani, N., Ferrari, P., Norat, T., Fahey, M., Charrondiere, U.R., Hemon, B., Casagrande, C., Vignat, J., et al. (2002). European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public health nutrition 5, 1113-1124.
5. Rimm, E.B., Giovannucci, E.L., Willett, W.C., Colditz, G.A., Ascherio, A., Rosner, B., and Stampfer, M.J. (1991). Prospective study of alcohol consumption and risk of coronary disease in men. Lancet 338, 464-468.
6. Giles, G.G., and English, D.R. (2002). The Melbourne Collaborative Cohort Study. IARC scientific publications 156, 69-70.
7. Colditz, G.A., and Hankinson, S.E. (2005). The Nurses' Health Study: lifestyle and health among women. Nature reviews Cancer 5, 388-396.
8. Hankinson, S.E., Willett, W.C., Manson, J.E., Hunter, D.J., Colditz, G.A., Stampfer, M.J., Longcope, C., and Speizer, F.E. (1995). Alcohol, height, and adiposity in relation to estrogen and prolactin levels in postmenopausal women. Journal of the National Cancer Institute 87, 1297-1302.
9. Toniolo, P.G., Levitz, M., Zeleniuch-Jacquotte, A., Banerjee, S., Koenig, K.L., Shore, R.E., Strax, P., and Pasternack, B.S. (1995). A prospective study of endogenous estrogens and breast cancer in postmenopausal women. Journal of the National Cancer Institute 87, 190-197.
10. Gu, Y., Shore, R.E., Arslan, A.A., Koenig, K.L., Liu, M., Ibrahim, S., Lokshin, A.E., and Zeleniuch-Jacquotte, A. (2010). Circulating cytokines and risk of B-cell non-Hodgkin lymphoma: a prospective study. Cancer causes & control : CCC 21, 1323-1333.
11. Troy, J.D., Hartge, P., Weissfeld, J.L., Oken, M.M., Colditz, G.A., Mechanic, L.E., and Morton, L.M. (2010). Associations between anthropometry, cigarette smoking, alcohol consumption, and non-Hodgkin lymphoma in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. American journal of epidemiology 171, 1270-1281.
12. Hayes, R.B., Sigurdson, A., Moore, L., Peters, U., Huang, W.Y., Pinsky, P., Reding, D., Gelmann, E.P., Rothman, N., Pfeiffer, R.M., et al. (2005). Methods for etiologic and early marker investigations in the PLCO trial. Mutation research 592, 147-154.
13. Anderson, G.L., Manson, J., Wallace, R., Lund, B., Hall, D., Davis, S., Shumaker, S., Wang, C.Y., Stein, E., and Prentice, R.L. (2003). Implementation of the Women's Health Initiative study design. Annals of epidemiology 13, S5-17.
14. Spinelli, J.J., Ng, C.H., Weber, J.P., Connors, J.M., Gascoyne, R.D., Lai, A.S., Brooks-Wilson, A.R., Le, N.D., Berry, B.R., and Gallagher, R.P. (2007). Organochlorines and risk of non-Hodgkin lymphoma. International journal of cancer Journal international du cancer 121, 2767-2775.
15. Willett, E.V., Smith, A.G., Dovey, G.J., Morgan, G.J., Parker, J., and Roman, E. (2004). Tobacco and alcohol consumption and the risk of non-Hodgkin lymphoma. Cancer causes & control : CCC 15, 771-780.

16. Worrillow, L., Roman, E., Adamson, P.J., Kane, E., Allan, J.M., and Lightfoot, T.J. (2009). Polymorphisms in the nucleotide excision repair gene ERCC2/XPD and risk of non-Hodgkin lymphoma. Cancer epidemiology 33, 257-260.
17. Crouch, S., Simpson, J., Ansell, P., Kane, E., Howell, D., Smith, A., Newton, R., Jack, A., and Roman, E. (2011). Illness patterns prior to diagnosis of lymphoma: analysis of UK medical records. Cancer epidemiology 35, 145-150.
18. Chatterjee, N., Hartge, P., Cerhan, J.R., Cozen, W., Davis, S., Ishibe, N., Colt, J., Goldin, L., and Severson, R.K. (2004). Risk of non-Hodgkin's lymphoma and family history of lymphatic, hematologic, and other cancers. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 13, 1415-1421.
19. Wang, S.S., Cerhan, J.R., Hartge, P., Davis, S., Cozen, W., Severson, R.K., Chatterjee, N., Yeager, M., Chanock, S.J., and Rothman, N. (2006). Common genetic variants in proinflammatory and other immunoregulatory genes and risk for non-Hodgkin lymphoma. Cancer research 66, 9771-9780.
20. Hughes, A.M., Armstrong, B.K., Vajdic, C.M., Turner, J., Grulich, A., Fritschi, L., Milliken, S., Kaldor, J., Benke, G., and Kricker, A. (2004). Pigmentary characteristics, sun sensitivity and non-Hodgkin lymphoma. International journal of cancer Journal international du cancer 110, 429-434.
21. Smedby, K.E., Hjalgrim, H., Melbye, M., Torrang, A., Rostgaard, K., Munksgaard, L., Adami, J., Hansen, M., Porwit-MacDonald, A., Jensen, B.A., et al. (2005). Ultraviolet radiation exposure and risk of malignant lymphomas. Journal of the National Cancer Institute 97, 199-209.
22. Skibola, C.F., Bracci, P.M., Halperin, E., Nieters, A., Hubbard, A., Paynter, R.A., Skibola, D.R., Agana, L., Becker, N., Tressler, P., et al. (2008). Polymorphisms in the estrogen receptor 1 and vitamin C and matrix metalloproteinase gene families are associated with susceptibility to lymphoma. PloS one 3, e2816.
23. Zhang, Y., Hughes, K.J., Zahm, S.H., Zhang, Y., Holford, T.R., Dai, L., Bai, Y., Han, X., Qin, Q., Lan, Q., et al. (2009). Genetic variations in xenobiotic metabolic pathway genes, personal hair dye use, and risk of non-Hodgkin lymphoma. American journal of epidemiology 170, 1222-1230.
24. Monnereau, A., Orsi, L., Troussard, X., Berthou, C., Fenaux, P., Soubeyran, P., Marit, G., Huguet, F., Milpied, N., Leporrier, M., et al. (2008). Cigarette smoking, alcohol drinking, and risk of lymphoid neoplasms: results of a French case-control study. Cancer causes & control : CCC 19, 1147-1160.
25. Besson, H., Brennan, P., Becker, N., Nieters, A., De Sanjose, S., Font, R., Maynadie, M., Foretova, L., Cocco, P.L., Staines, A., et al. (2006). Tobacco smoking, alcohol drinking and non-Hodgkin's lymphoma: A European multicenter case-control study (Epilymph). International journal of cancer Journal international du cancer 119, 901-908.
26. Drake, M.T., Maurer, M.J., Link, B.K., Habermann, T.M., Ansell, S.M., Micallef, I.N., Kelly, J.L., Macon, W.R., Nowakowski, G.S., Inwards, D.J., et al. (2010). Vitamin D insufficiency and prognosis in non-Hodgkin's lymphoma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 28, 4191-4198.
27. Cerhan, J.R., Fredericksen, Z.S., Wang, A.H., Habermann, T.M., Kay, N.E., Macon, W.R., Cunningham, J.M., Shanafelt, T.D., Ansell, S.M., Call, T.G., et al. (2011). Design and validity of a clinic-based case-control study on the molecular epidemiology of lymphoma. International journal of molecular epidemiology and genetics 2, 95-113.
28. Yossepowitch, O., Olvera, N., Satagopan, J.M., Huang, H., Jhanwar, S., Rapaport, B., Boyd, J., and Offit, K. (2003). BRCA1 and BRCA2 germline mutations in lymphoma patients. Leukemia & lymphoma 44, 127-131.
29. Vijai, J., Kirchhoff, T., Schrader, K.A., Brown, J., Dutra-Clarke, A.V., Manschreck, C., Hansen, N., Rau-Murthy, R., Sarrel, K., Przybylo, J., et al. (2013). Susceptibility loci associated with specific and shared subtypes of lymphoid malignancies. PLoS genetics 9, e1003220.

30. Salles, G., Seymour, J.F., Offner, F., Lopez-Guillermo, A., Belada, D., Xerri, L., Feugier, P., Bouabdallah, R., Catalano, J.V., Brice, P., et al. (2011). Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomised controlled trial. Lancet 377, 42-51.
31. De Vivo, I., Prescott, J., Setiawan, V.W., Olson, S.H., Wentzensen, N., Australian National Endometrial Cancer Study, G., Attia, J., Black, A., Brinton, L., Chen, C., et al. (2014). Genome-wide association study of endometrial cancer in E2C2. Human genetics 133, 211-224.
32. Holly, E.A., and Bracci, P.M. (2003). Population-based study of non-Hodgkin lymphoma, histology, and medical history among human immunodeficiency virus-negative participants in San Francisco. American journal of epidemiology 158, 316-327.
33. Morton, L.M., Turner, J.J., Cerhan, J.R., Linet, M.S., Treseler, P.A., Clarke, C.A., Jack, A., Cozen, W., Maynadie, M., Spinelli, J.J., et al. (2007). Proposed classification of lymphoid neoplasms for epidemiologic research from the Pathology Working Group of the International Lymphoma Epidemiology Consortium (InterLymph). Blood 110, 695-708.
34. Turner, J.J., Morton, L.M., Linet, M.S., Clarke, C.A., Kadin, M.E., Vajdic, C.M., Monnereau, A., Maynadie, M., Chiu, B.C., Marcos-Gragera, R., et al. (2010). InterLymph hierarchical classification of lymphoid neoplasms for epidemiologic research based on the WHO classification (2008): update and future directions. Blood 116, e90-98.
35. Ansel, K.M., Ngo, V.N., Hyman, P.L., Luther, S.A., Forster, R., Sedgwick, J.D., Browning, J.L., Lipp, M., and Cyster, J.G. (2000). A chemokine-driven positive feedback loop organizes lymphoid follicles. Nature 406, 309-314.
36. Saez de Guinoa, J., Barrio, L., Mellado, M., and Carrasco, Y.R. (2011). CXCL13/CXCR5 signaling enhances BCR-triggered B-cell activation by shaping cell dynamics. Blood 118, 1560-1569.
37. Charbonneau, B., Wang, A.H., Maurer, M.J., Asmann, Y.W., Zent, C.S., Link, B.K., Ansell, S.M., Weiner, G.J., Ozsan, N., Feldman, A.L., et al. (2013). CXCR5 polymorphisms in non-Hodgkin lymphoma risk and prognosis. Cancer immunology, immunotherapy : CII 62, 1475-1484.
38. Russell, L., and Garrett-Sinha, L.A. (2010). Transcription factor Ets-1 in cytokine and chemokine gene regulation. Cytokine 51, 217-226.
39. Grunewald, T.G., Pasedag, S.M., and Butt, E. (2009). Cell Adhesion and Transcriptional Activity - Defining the Role of the Novel Protooncogene LPP. Translational oncology 2, 107-116.
40. Tan, D.E., Foo, J.N., Bei, J.X., Chang, J., Peng, R., Zheng, X., Wei, L., Huang, Y., Lim, W.Y., Li, J., et al. (2013). Genome-wide association study of B cell non-Hodgkin lymphoma identifies 3q27 as a susceptibility locus in the Chinese population. Nature genetics 45, 804-807.
41. Yunis, J.J., Frizzera, G., Oken, M.M., McKenna, J., Theologides, A., and Arnesen, M. (1987). Multiple recurrent genomic defects in follicular lymphoma. A possible model for cancer. The New England journal of medicine 316, 79-84.
42. Berndt, S.I., Skibola, C.F., Joseph, V., Camp, N.J., Nieters, A., Wang, Z., Cozen, W., Monnereau, A., Wang, S.S., Kelly, R.S., et al. (2013). Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. Nature genetics 45, 868-876.
43. Graham, M., and Adams, J.M. (1986). Chromosome 8 breakpoint far 3' of the c-myc oncogene in a Burkitt's lymphoma 2;8 variant translocation is equivalent to the murine pvt-1 locus. The EMBO journal 5, 2845-2851.
44. Tsutsumi, Y., Chinen, Y., Sakamoto, N., Nagoshi, H., Nishida, K., Kobayashi, S., Yokokawa, Y., Taki, T., Sasaki, N., Yamamoto-Sugitani, M., et al. (2013). Deletion or methylation of CDKN2A/2B and PVT1 rearrangement occur frequently in highly aggressive B-cell lymphomas harboring 8q24 abnormality. Leukemia & lymphoma 54, 2760-2764.
45. Enciso-Mora, V., Broderick, P., Ma, Y., Jarrett, R.F., Hjalgrim, H., Hemminki, K., van den Berg, A., Olver, B., Lloyd, A., Dobbins, S.E., et al. (2010). A genome-wide association study of Hodgkin's

lymphoma identifies new susceptibility loci at 2p16.1 (REL), 8q24.21 and 10p14 (GATA3). Nature genetics 42, 1126-1130.
46. Crowther-Swanepoel, D., and Houlston, R.S. (2010). Genetic variation and risk of chronic lymphocytic leukaemia. Seminars in cancer biology 20, 363-369.
47. Gudmundsson, J., Sulem, P., Gudbjartsson, D.F., Blondal, T., Gylfason, A., Agnarsson, B.A., Benediktsdottir, K.R., Magnusdottir, D.N., Orlygsdottir, G., Jakobsdottir, M., et al. (2009). Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. Nature genetics 41, 1122-1126.
48. Rothman, N., Garcia-Closas, M., Chatterjee, N., Malats, N., Wu, X., Figueroa, J.D., Real, F.X., Van Den Berg, D., Matullo, G., Baris, D., et al. (2010). A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. Nature genetics 42, 978-984.
49. Michailidou, K., Hall, P., Gonzalez-Neira, A., Ghoussaini, M., Dennis, J., Milne, R.L., Schmidt, M.K., Chang-Claude, J., Bojesen, S.E., Bolla, M.K., et al. (2013). Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nature genetics 45, 353-361, 361e351-352.
50. Zanke, B.W., Greenwood, C.M., Rangrej, J., Kustra, R., Tenesa, A., Farrington, S.M., Prendergast, J., Olschwang, S., Chiang, T., Crowdy, E., et al. (2007). Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nature genetics 39, 989-994.
51. Tomlinson, I., Webb, E., Carvajal-Carmona, L., Broderick, P., Kemp, Z., Spain, S., Penegar, S., Chandler, I., Gorman, M., Wood, W., et al. (2007). A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nature genetics 39, 984-988.
52. Skibola, C.F., Bracci, P.M., Halperin, E., Conde, L., Craig, D.W., Agana, L., Iyadurai, K., Becker, N., Brooks-Wilson, A., Curry, J.D., et al. (2009). Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. Nature genetics 41, 873-875.
53. Nieters, A., Conde, L., Slager, S.L., Brooks-Wilson, A., Morton, L., Skibola, D.R., Novak, A.J., Riby, J., Ansell, S.M., Halperin, E., et al. (2012). PRRC2A and BCL2L11 gene variants influence risk of non-Hodgkin lymphoma: results from the InterLymph consortium. Blood 120, 4645-4648.
54. Smedby, K.E., Foo, J.N., Skibola, C.F., Darabi, H., Conde, L., Hjalgrim, H., Kumar, V., Chang, E.T., Rothman, N., Cerhan, J.R., et al. (2011). GWAS of follicular lymphoma reveals allelic heterogeneity at 6p21.32 and suggests shared genetic susceptibility with diffuse large B-cell lymphoma. PLoS genetics 7, e1001378.
55. Conde, L., Halperin, E., Akers, N.K., Brown, K.M., Smedby, K.E., Rothman, N., Nieters, A., Slager, S.L., Brooks-Wilson, A., Agana, L., et al. (2010). Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. Nature genetics 42, 661-664.
56. Cerhan, J.R., Fredericksen, Z.S., Novak, A.J., Ansell, S.M., Kay, N.E., Liebow, M., Dogan, A., Cunningham, J.M., Wang, A.H., Witzig, T.E., et al. (2012). A two-stage evaluation of genetic variation in immune and inflammation genes with risk of non-Hodgkin lymphoma identifies new susceptibility locus in 6p21.3 region. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 21, 1799-1806.
57. Wang, S.S., Abdou, A.M., Morton, L.M., Thomas, R., Cerhan, J.R., Gao, X., Cozen, W., Rothman, N., Davis, S., Severson, R.K., et al. (2010). Human leukocyte antigen class I and II alleles in non-Hodgkin lymphoma etiology. Blood 115, 4820-4823.
58. Skibola, C.F., Akers, N.K., Conde, L., Ladner, M., Hawbecker, S.K., Cohen, F., Ribas, F., Erlich, H.A., Goodridge, D., Trachtenberg, E.A., et al. (2012). Multi-locus HLA class I and II allele and haplotype associations with follicular lymphoma. Tissue antigens 79, 279-286.
59. Foo, J.N., Smedby, K.E., Akers, N.K., Berglund, M., Irwan, I.D., Jia, X., Li, Y., Conde, L., Darabi, H., Bracci, P.M., et al. (2013). Coding variants at hexa-allelic amino acid 13 of HLA-DRB1 explain

independent SNP associations with follicular lymphoma risk. American journal of human genetics 93, 167-172.