HLA Class I and II Diversity Contributes to the Etiologic Heterogeneity of Non-Hodgkin Lymphoma Subtypes

Sophia S. Wang, Beckman Research Institute
Mary Carrington, Leidos Biomedical Research Inc.
Sonja I. Berndt, National Cancer Institute
Susan L. Slager, Mayo Clinic
Paige M. Bracci, University of California San Francisco
Jenna Voutsinas, Beckman Research Institute
James R. Cerhan, Mayo Clinic
Karin E. Smedby, Karolinska Institutet
Henrik Hjalgrim, Statens Serum Institut
Joseph Vijai, Memorial Sloan Kettering Cancer Center

Only first 10 authors above; see publication for full author list.

Journal Title: Cancer Research
Volume: Volume 78, Number 14
Publisher: American Association for Cancer Research | 2018-07-15, Pages 4086-4096
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1158/0008-5472.CAN-17-2900
Permanent URL: https://pid.emory.edu/ark:/25593/txsvh

Final published version: http://dx.doi.org/10.1158/0008-5472.CAN-17-2900

Copyright information:
© 2018 American Association for Cancer Research.

Accessed April 25, 2020 12:27 AM EDT
HLA class I and II diversity contributes to the etiologic heterogeneity of non-Hodgkin lymphoma subtypes

Abstract

A growing number of loci within the human leukocyte antigen (HLA) region have been implicated in non-Hodgkin lymphoma (NHL) etiology. Here, we test a complementary hypothesis of “heterozygote advantage” regarding the role of HLA and NHL, whereby HLA diversity is beneficial and homozygous HLA loci are associated with increased disease risk. HLA alleles at class I and II loci were imputed from genome-wide association studies (GWAS) using SNP2HLA for: 3,617 diffuse large B-cell lymphomas (DLBCL), 2,686 follicular lymphomas (FL), 2,878 chronic lymphocytic leukemia/small lymphocytic lymphomas (CLL/SLL), 741 marginal zone lymphomas (MZL), and 8,753 controls of European descent. Both DLBCL and MZL risk were elevated with homozygosity at class I HLA-B and -C loci (OR DLBCL=1.31, 95% CI=1.06–1.60; OR MZL=1.45, 95% CI=1.12–1.89) and class II HLA-DRB1 locus (OR DLBCL=2.10, 95% CI=1.24–3.55; OR MZL=2.10, 95% CI=0.99–4.45). Increased FL risk was observed with the overall increase in number of homozygous HLA class II loci (p-trend<0.0001, FDR=0.0005). These results support a role for HLA zygosity in NHL etiology and suggests that distinct immune pathways may underly the etiology of the different NHL subtypes.

INTRODUCTION

Genome-wide association studies (GWAS) have identified a growing list of common susceptibility loci modestly associated with risk of non-Hodgkin lymphomas (NHLs) including several HLA (human leukocyte antigen) genetic variants on chromosome 6p21, a region that is critical for innate and adaptive immune responses. Putative NHL susceptibility loci either directly implicate genes within the Major Histocompatibility Complex (MHC) or appear in strong linkage disequilibrium (LD) with extended HLA haplotypes (1–5). Interestingly, there is little convincing overlap of the identified HLA susceptibility loci among the NHL subtypes, diffuse large B-cell lymphoma (DLBCL), follicular lymphoma (FL), marginal zone lymphoma (MZL) and chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL), suggesting that disparate aspects of the MHC and resulting immune responses are involved in the etiology of each NHL subtype.

The HLA genes are the most polymorphic in the human genome and specific HLA loci determine the antigens that are bound by antigen presenting cells (e.g., B cells and dendritic cells) and presented to T cells to elicit immune responses. Functionally, HLA molecules are
critical for the host immune response. HLA class I molecules present foreign antigens primarily to cytotoxic T-cells that in response kill these target cells, while HLA class II molecules stimulate antibody production in response to specific antigens.

Reduced diversity, as defined by homozygosity at each co-dominant HLA loci, might adversely affect the host’s ability to recognize a more diverse array of foreign antigens and thereby increase subsequent disease burden. This concept is supported by a priori research that has examined effects of HLA zygosities on infectious disease, whereby a lack of HLA class I and II diversity has been associated with increased risk HIV and hepatitis B virus infection (6–8).

Given the growing evidence that genetic variation within HLA genes play in the etiology of NHL subtypes (1–4, 9), we specifically aimed to test whether lack of HLA diversity - as measured by HLA homozygosity – was associated with increased NHL risk. Specifically, we posit that associations with HLA Class II, which primarily presents peptides derived from extracellular sources, would implicate a role in infectious disease etiology. On the other hand, associations with HLA Class I, which primarily presents peptides derived from intracellular sources, would suggest a role in related conditions, such as autoimmune or atopic conditions. We present here results from a pooled analysis of 25 studies from North America, Europe, and Australia where we measured the associations between HLA class I and/or class II zygosities and four main NHL subtypes.

MATERIALS AND METHODS

Study sample

Our study sample comprises the same study participants of European descent that were included in the original GWAS efforts from which 25 studies participated. Specifically, adults diagnosed with incident, non-HIV-related B-cell NHL of mostly European descent, ascertained from cancer registries, clinics, or hospitals or through self-report were included and where diagnoses were verified by medical and pathology reports (1–4). Study designs included prospective cohort studies, population- and hospital-based case-control studies, and clinic-based studies. Original details of design methods for each study and of each GWAS have been described previously (1–4).

This study was approved by the City of Hope Institutional Review Board. Each participating study obtained approval from human subjects review committees and written informed consent from all participants. A de-identified pooled dataset with individual-level data on genotypes, demographic characteristics, and NHL subtypes of cases was provided by the InterLymph Data Coordinating Center (Mayo Clinic, Rochester, MN).

Genotyping

GWAS platforms used include the Illumina 317K, Illumina HumanHap 610K, Illumina HumanHap 660W, Illumina Human CNV370-Duo BeadChip, Affymetrix SNP 6.0, and the Illumina OmniExpress (Table 1). Quality control metrics employed (e.g., QQ plots and Eigenstrat results) and main results of each GWAS have been previously described in-depth (1–4).
**HLA imputation**

As reported by Skibola et al (2), classical HLA alleles were imputed at HLA class I (HLA-A, HLA-B, HLA-C) and class II loci (HLA-DQA1, HLA-DQB1, HLA-DRB1, HLA-DPA1, HLA-DPB1) using SNP2HLA and a reference panel from the Type 1 Diabetes Genetics Consortium that comprised 5,225 individuals of European descent who were typed for HLA-A, B, C, DQA1, DQB1, DRB1, DPA1, DPB1 4 digit alleles. We note that the SNP2HLA reference panel is typed both for a panel of MHC SNPs and using classical HLA typing; the imputation algorithms used thus rely on both methodologies particularly when only SNPs are available. A comparison of imputed HLA alleles to 4-digit HLA sequencing data available for a subset of samples showed high concordance: HLA-A (97.3%), B (98.5%), C (98.1%) and DRB1 (97.5%). In all, 201 classical HLA alleles (two- and four-digit resolution) were successfully imputed (info score \( r^2 > 0.3 \) for alleles) and available for analysis. Because of the strong LD between the HLA class II A1 and B1 loci (e.g., HLA-DQA1 and DQB1), we present results for each of the B1 loci (HLA-DQB1, HLA-DRB1, HLA-DPB1) since there were fewer homozygous B1 loci than A1 loci. For each HLA locus, individuals were coded as homozygote (for any allele) or heterozygote, as determined from the imputed alleles. All results presented are based on four-digit resolution.

**NHL Classification**

NHL subtypes were harmonized at the InterLymph Data Coordinating Center using the InterLymph Pathology Working Group guidelines (10,11), which are based on the World Health Organization classification (12).

**Final analytic sample**

Data for HLA loci were directly imputed from the original GWAS SNP panels and evaluated for the 3,617 DLBCL, 2,686 FL, 2,878 CLL/SLL, 741 MZL, and 8,753 controls. We note that, as with the original GWAS manuscripts, the specific numbers of controls differed by NHL subtype, due to different study inclusion and control selection criteria for each NHL subtype analyses, as described by the original GWAS publications (enumerated in Table 2).

**Statistical analysis**

Heterozygosity and homozygosity at each individual HLA locus and the number of homozygous loci for class I loci (A, B, C) and class II loci (DQB1, DRB1, DPB1) were determined; odds ratios (ORs) and 95% confidence intervals (CIs) were calculated as estimates of NHL risk with heterozygotes as the referent category, adjusted for sex, age, study, GWAS platform, and ancestry (with principal components as conducted for each subtype-specific GWAS and previously published (1–4). For analyses of MZL, adjustment by geographic region was conducted due to sample size restrictions (instead of by individual study). In addition to calculating the risk estimates for each additional number of homozygous loci, we further calculated the p-trend.

To further describe associations of zygosity by loci, we conducted joint effects analyses for HLA class I loci and class II loci. Each HLA loci (class I or II) was conducted in a stratified manner whereby heterozygotes for all loci were the referent groups and all combinations of homozygosity among the loci were evaluated. For example, to pinpoint whether HLA class I
associations were attributable to HLA Class I B or C loci, we modeled as one covariate, 4 levels/combinations for HLA-B and -C (e.g., homozygous for both HLA-B and -C, homozygous only for HLA-B, homozygous for only HLA-C, and heterozygous for both), with heterozygote for both HLA-B and -C as reference (Table 3). For the associated p-trends reported in Table 3, each category is modeled based on ordinal variable in the order listed in the table, with heterozygosity at all loci as the referent group in a logistic regression model. For each p-trend, we also present the linearized additive relative-risk-per-locus, reflecting the slope of the trend-line.

Platform-specific results are shown in a Supplemental Table 1. Additional sensitivity analysis included evaluation of potential confounders, including evaluation of associations by previously implicated autoimmune conditions and HLA loci associated with specific NHL subtypes. We conducted stratified analysis to evaluate whether HLA zygosity associations were present among participants with and without autoimmune conditions (generally, and by specific conditions); similarly stratified analyses were conducted among participants with and without previously identified SNPs associated with NHL subtypes. We further calculated the risks, adjusting for autoimmune conditions and for all reported genetic susceptibility loci (for each NHL subtype). As neither variable altered the odds ratio >10%, those data are not presented. Analyses that restricted studies to population-based controls only also did not have measurable effect on the results. Finally, to evaluate the probability that some of our results could be due to chance, we used the Benjamini-Hochberg method to calculate the false discovery rate (FDR) and applied it to the p-trends as this allows for the fewest number of comparisons and thus degrees of freedom to assess the additive model.

Unconditional logistic regression models were applied using SAS 9.4 (SAS Institute). All tests of statistical significance were 2-sided.

RESULTS

The numbers of European cases and controls from each of the 25 studies in North America, Europe, and Australia for which HLA class I and II loci were evaluated are detailed in Table 1.

DLBCL

Elevated DLBCL risks of 20–50% were observed for homozygosity for individual HLA class I (B and C) and/or class II loci (DRB1 and DQB1) (Table 2). DLBCL risk also increased with increasing number of homozygous class I loci (p-trend=0.0008; FDR p=0.003) and class II loci (p-trend<0.0001; FDR p=0.0005) (Table 2). Although homozygosity for HLA-A had a borderline non-significant effect for increasing DLBCL risk, joint analyses suggested that the 30% risk increase observed with two or more homozygote loci (Table 2) was attributable to homozygosity at the HLA-B and -C locus (OR=1.31, 95% CI=1.06–1.60, Table 3). Similarly, for class II loci, joint analysis showed statistically significant associations for homozygosity specifically at the HLA-DRB1 locus (OR=2.10, 95% CI=1.24–3.55) as significantly increased risk was observed only in combination with homozygous HLA-DRB1 locus (Table 3).
FL

There were no significant associations between zygosity at *HLA* class I loci and FL risk (Table 2). Statistically significant 24–54% increases, however, were observed for FL risk for each of the three *HLA* class II loci. Further, FL risk increased with the total number of homozygous *HLA* class II loci (p-trend<0.0001; FDR p=0.0005), with an odds ratio of 1.89 (95% CI=1.37–2.61) for those fully homozygous compared with those fully heterozygous at all three *HLA* class II loci. Joint analyses additionally supported a statistically significant increased risk for FL with overall homozygosity at the *HLA* class II loci (p-trend<0.0001; FDR p=0.0005, Table 3).

MZL

Homozygosity at *HLA* class I loci *HLA*-B (OR=1.34, 95% CI=1.01–1.78) and –C (OR=1.33, 95% CI=1.04–1.70) but not –A (OR=1.06, 95% CI=0.82–1.38) increased MZL risk (Table 2). Stratified analysis supported independent associations for both *HLA*-B and –C and MZL (Table 3). Homozygosity at *HLA* class II loci increased MZL risk (Table 2), but only the association with *HLA-DRB1* reached statistical significance (OR=1.45, 95% CI=1.12–1.89, Table 2). Analyses considering single locus homozygosity provided evidence of a role for *HLA-DRB1* in increasing MZL risk (Table 3).

CLL/SLL

Modest CLL/SLL risk increases were observed for *HLA*-A (OR=1.19, 95% CI=1.02–1.38), *HLA-DRB1* (OR=1.19, 95% CI=1.00–1.42) and *HLA-DQB1* (OR=1.20, 95% CI=1.03–1.39) (Table 2). Increasing CLL/SLL risk was not observed with increasing number of homozygote class I or class II loci, though when evaluating total numbers of class I and II loci altogether, a borderline significant increased risk was observed for those with all five homozygote class I and II loci (OR=1.57, 95% CI=1.04–2.38, p-trend = 0.029; FDR=0.055) (Table 2). We were unable to isolate CLL/SLL associations with *HLA* zygosity to any singular locus (Table 3).

DISCUSSION

Based on the largest number of NHL subtypes to date for whom imputed HLA data is available, we demonstrate that *HLA* homozygosity plays a role in four B-cell NHL subtypes, and that the associations between homozygosity at *HLA* Class I and/or Class II loci are distinct by these subtypes. Specifically, FL risk was associated with homozygosity at *HLA* class II loci, but not Class I loci. CLL/SLL risk appeared to be associated (borderline) with homozygosity at either *HLA* Class I or Class II loci. In contrast, while both DLBCL and MZL were associated with zygosity at *HLA* Class I and Class II loci, the associations appeared specific to Class I *HLA*-B and –C loci and to the Class II *HLA-DRB1* locus. We note that the p-trends evaluated for each additional homozygous loci remained statistically significant after adjust for multiple comparisons, with exception of that for CLL/SLL. Our results add to the growing body of literature implicating different roles for *HLA* class I and II loci, key modulators of human immune response, in the heterogeneous etiologies of B-NHL subtypes (1–4). Our results also add to the current literature which points to similarities in the etiologic profiles of DLBCL and MZL (13). Overall, these data support
the importance of HLA diversity in NHL etiology, with the type of HLA diversity potentially varying by NHL subtype.

The underlying hypothesis regarding the role of HLA zygosity and disease is that homozygosity at HLA loci reduces the diversity of peptides that can be presented, with the hypothesis that these peptides can reflect etiologic agents such as infectious diseases, self-antigens for atopic or autoimmune conditions, and even cancerous cells. At present, there is a growing body of literature supporting that HLA heterozygotes are more resistant to infectious diseases, and the corollary, that HLA homozygotes are more susceptible to infectious diseases. Specifically, HLA class I heterozygote advantage (e.g., presenting greater diversity of antigenic peptides to CD8+ cytotoxic T lymphocytes) has been demonstrated for slowing progression to AIDS (6), whereas heterozygotes at HLA class II loci appear to have greater ability in clearing HBV infection (8) and HCV infection (14) than homozygotes. HLA-DRB1 heterozygosity has also been reported to confer favorable outcome (e.g., against end-stage liver disease) among HCV-infected liver transplant recipients (15). There are also reports evaluating HLA zygosity as a key contributor in autoimmune conditions. For example, reports of heterozygote advantage for class II loci and inflammatory bowel disease (16) and for class I loci and psoriatic arthritis (17) have both been published. Specific associations between HLA zygosity and NHL have been limited to reports of CLL. Evidence of the importance of HLA zygosity include reports that homozygosity at HLA-A, -B, and -DRB1 are associated with CLL (18) and with CLL disease progression (19–20), with the hypothesis that limited HLA diversity provided an advantage of the tumor to escape the immune response.

HLA heterozygote advantage is posited to work in concert with specific allele associations (as opposed to exclusively) (21); our results thus complement ongoing efforts that have identified the most role that specific HLA alleles have on NHL subtype risk. In sensitivity analysis, we evaluated the effect of known HLA associations and, in stratified and adjusted analysis, did not find that these associations diminish the reported association between HLA zygosity and NHL subtypes. Further evaluation into how these complementary associations act in concert are thus warranted and inclusion of HLA zygosity in the construct of genetic risk scores for each NHL subtype should be considered.

Further research to understand the association – or independence - between HLA zygosity with infections and autoimmune conditions and NHL risk are also needed (21–24). For example, efforts to evaluate autoimmune conditions linked to class II alleles (e.g., Sjögren syndrome, systemic lupus erythematosus, and rheumatoid arthritis) (23) with class II zygosity in relation to FL risk could provide potential insight regarding immune mechanisms modulating FL risk. A particularly pressing research question is understanding what are the underlying mechanisms of individual allele-associations and how are they distinct from HLA zygosity associations. Similar efforts to identify commonalities between autoimmune and infectious disease associations with HLA loci and zygosity among other NHL subtypes are also warranted. Finally, extension of these efforts towards understanding the genetic and structural variants and HLA expression are also required to fully understand the implication of HLA-allelic associations in the context of overall class I or II zygosity.
Study strengths include the large sample size available to evaluate individual NHL subtypes which no studies have been able to do adequately to date (25). Potential study limitations include possible misclassification of HLA alleles due to imputation, although direct comparison of a subset with genotyped HLA alleles showed >97% concordance (2). While the present analysis leverages the available GWAS data through imputation of HLA alleles, we recognize that confirmation with direct HLA allelotyping may provide additional levels of information not ascertained in imputed data.

Our study’s restriction to individuals of European ancestry requires our results to be replicated for other racial or ethnic groups, as the associations may not apply universally to all ethnic groups. However, as demonstrated for HLA associations in autoimmune conditions, fine-mapping studies show that the same amino acid changes contribute to disease in both European and Asian populations (26), implicating similar underlying biologic mechanisms for disease etiology. Studies limitations also include our inability to evaluate heterogeneity within NHL subtypes, either defined molecularly, by infectious etiology, or by organ site.

In summary, our results add to the growing evidence of HLA alleles as susceptibility loci in the etiology of B-cell NHL subtypes. In addition to ongoing fine-mapping studies being conducted as follow-up to GWAS, our results here suggest that functional studies aiming to understand the underlying biology of zygosity and NHL subtype risk will also be important. Additional efforts to evaluate larger-scale zygosity, such as of immune genes and perhaps the entire genome may prove important in understanding the full extent of the role diversity of the immune response plays in lymphoma etiology.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Authors**

Karen Curtin54, Jacqueline Clavel55,56, Alain Monnereau55,56,57, David G. Cox58, Hervé Ghesquières59,60, Gilles Salles60,61, Paulo Bofetta62, Lenka Foretova63, Anthony Staines64, Scott Davis65, Richard K. Severson66, Qing Lan3, Angela Brooks-Wilson67,68, Martyn T Smith69, Eve Roman20, Anne Kricker70, Yawei Zhang71,72, Peter Kraft73, Stephen J. Chanock3, Nathaniel Rothman3, Patricia Hartge3, and Christine F. Skibola27

Affiliations

1Department of Population Sciences, Beckman Research Institute and the City of Hope, Duarte, California 2Cancer and Inflammation Program, Leidos Biomedical Research Inc., Frederick National Laboratory for Cancer Research, Frederick, MD and Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA 3Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD 4Department of Health Sciences Research, Mayo Clinic, Rochester, MN 5Department of Epidemiology and Biostatistics, University of California–San Francisco, San Francisco, CA 6Department of Medicine Solna, unit of clinical epidemiology, Karolinska Institutet, Stockholm, Sweden 7Hematology Center, Karolinska University Hospital, Stockholm, Sweden 8Department of Epidemiology Research, Division of Health Surveillance and Research, Statens Serum Institut, Copenhagen, Denmark 9Department of Hematology, Rishospitalet, Copenhagen, Denmark 10Department of Medicine, Memorial Sloan Kettering Cancer Center, New York City, NY 11Institute for Risk Assessment Sciences, Utrecht University, Utrecht, the Netherlands 12Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands 13Braun School of Public Health and Community Medicine, Hadassah-Hebrew University Medical Center, Jerusalem, Israel 14Centre for Big Data Research in Health, The University of New South Wales, Sydney, New South Wales, Australia 15Centre for Chronic Immunodeficiency, University Medical Center Freiburg, Freiburg, Baden-Württemberg, Germany 16Unit of Infections and Cancer, Cancer Epidemiology Research Programme, Institut Català d’Oncologia, IDIBELL, 08908 L’Hospitalet de Llobregat, Barcelona, Spain 17CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain 18Norris Comprehensive Cancer Center, Keck School of Medicine of USC, Departments of Preventive Medicine and Pathology, University of Southern California, Los Angeles, CA 19Department of Pathology, School of Medicine and the UAB Comprehensive Cancer Center, The University of Alabama at Birmingham, Birmingham, AL 20Department of Health Sciences, University of York, York, United Kingdom 21Department of Histopathology, Douglass Hanly Moir Pathology, Sydney, Australia 22Faculty of Medicine and Health Sciences, Macquarie University, Sydney, Australia 23Cancer Control Research, British Columbia Cancer Agency, Vancouver, British Columbia, Canada 24School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada 25Department of Epidemiology, School of Public Health, Brown University, Providence, RI 26Channing Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA 27Department of Hematology and Medical Oncology, Emory University School of Medicine, Atlanta, GA 28Division of Clinical Epidemiology, German Cancer
Research Centre, Heidelberg, Baden-Württemberg, Germany  
Department of Pathology, City of Hope, Duarte, California  
Registry of Hematological Malignancies of Cote d’Or, INSERM UMR1231, University of Burgundy and Dijon University Hospital, Dijon, France  
Department of Public Health, Clinical and Molecular Medicine, University of Cagliari, Cagliari, Italy  
Epidemiology Research Program, American Cancer Society, Atlanta, GA  
Cancer Epidemiology Unit, University of Oxford, Oxford, United Kingdom  
School of Public Health, Imperial College London, London, United Kingdom  
International Agency for Research on Cancer, Lyon, France  
Department of Immunology, CHU Henri Mondor, Créteil, France  
Department of Internal Medicine, Carver College of Medicine, The University of Iowa, Iowa City, IA  
Center of Oncological Prevention (CPO) Piemonte and Unit of Medical Statistics and Epidemiology, Department Translational Medicine, University of Piemonte Orientale, Novara, Italy  
Department of Biomedical Science, University of Cagliari, Monserrato, Cagliari, Italy  
Division of Hematology, S. Francesco Hospital, Nuoro, Italy  
Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN  
Cancer Epidemiology & Intelligence Division, Cancer Council Victoria, Melbourne, Australia  
Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia  
Genetic Epidemiology Laboratory, Department of Pathology, University of Melbourne, Melbourne, Victoria, Australia  
Department of Obstetrics and Gynecology, New York University School of Medicine, New York City, NY  
Department of Environmental Medicine, New York University School of Medicine, New York City, NY  
Perlmutter Cancer Center, NYU Langone Medical Center, New York City, NY  
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden  
Department of Medicine, Stanford University School of Medicine, Stanford, CA  
Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden  
Bill Lyons Informatics Centre, UCL Cancer Institute, University College London, WC1E 6DD, London, United Kingdom  
Department of Internal Medicine, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT  
Epidemiology of Childhood and Adolescent Cancers Group, Inserm, Center of Research in Epidemiology and Statistics Sorbonne Paris Cité (CRESS), Paris, France  
Université Paris Descartes, Paris, France  
Registre des hémopathies malignes de la Gironde, Institut Bergonié, University of Bordeaux, Inserm, Team EPICENE, UMR 1219, France  
Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom. Cancer Research Center of Lyon, INSERM UMR1052, Center Léon Bérard, Lyon, France  
Department of Hematology, Centre Léon Bérard, Lyon, France  
Laboratoire de Biologie Moléculaire de la Cellule UMR 5239, Centre National de la Recherche Scientifique, Pierre benite Cedex, France  
Department of Hematology, Hospices Civils De Lyon, Centre Hospitalier Lyon-Sud and Université Claude Bernard, Lyon, France  
The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York City, New York  
Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, MF MU,
Acknowledgments

This analysis was initiated and conducted by the InterLymph Consortium Immunology and Infections Working Group and the Genome-Wide Association Study. We are grateful to the members of the working group, study investigators from contributing InterLymph case-control studies, and the InterLymph Consortium members for their contributions to this international collaboration. We are also thankful for the contributions from Aaron Norman and Dennis Robinson of the InterLymph Data Coordinating Center at the Mayo Clinic and to Michelle Dich at the City of Hope.

FUNDING ACKNOWLEDGEMENTS

This project was supported in part with funding from the National Institutes of Health (R01CA179558 and R01CA33572).

ATBC - The ATBC Study is supported by the Intramural Research Program of the U.S. National Cancer Institute, National Institutes of Health, and by U.S. Public Health Service contract HHSN261201500005C from the National Cancer Institute, Department of Health and Human Services.

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

CPS-II - The Cancer Prevention Study-II (CPS-II) Nutrition Cohort is supported by the American Cancer Society. Genotyping for all CPS-II samples were supported by the Intramural Research Program of the National Institutes of Health, NCI, Division of Cancer Epidemiology and Genetics. 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 – Bloodwise (formerly Leukaemia & Lymphoma Research), UK.

ENGELA – Association pour la Recherche contre le Cancer (ARC), Institut National du Cancer (INCa), Fondation de France, Fondation contre la Leucémie, Agence nationale de sécurité sanitaire de l’alimentation, de l’environnement et du travail (ANSES)

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

EpiLymph – European Commission (grant references QLK4-CT-2000-00422 and FOOD-CT-2006-023103); the Spanish Ministry of Health (grant references CIBERESP, PI11/01810, PI14/01219, RCESP C03/09, RTICESP C03/10 and RTIC RD06/0020/00095), the Marató de TV3 Foundation (grant reference 051210), the Agència de Gestió d’Ajuts Universitaris de Recerca – Generalitat de Catalunya (grant reference 2014SRG756) 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-R1223, 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 (MCCI, 00209805) and MEYS – NPS I – LO1413; Fondation de France and Association de Recherche Contre le Cancer.

FNLCR - This project has been funded in whole or in part with federal funds from the Frederick National Laboratory for Cancer Research, under Contract No. HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This Research was supported in part by the Intramural Research Program of the NIH, Frederick National Lab, Center for Cancer Research.

GEC/Mayo GWAS - National Institutes of Health (CA118444, CA148690, CA92153). Intramural Research Program of the NIH, National Cancer Institute. Veterans Affairs Research Service. Data collection for Duke University was supported by a Leukemia & Lymphoma Society Career Development Award, the Bernstein Family Fund for Leukemia and Lymphoma Research, and the National Institutes of Health (K08CA134919), National Center for Advancing Translational Science (UL1 TR000135).

HPFS – The HPFS was supported in part by National Institutes of Health grants CA167552, CA149445, and CA098122. We would like to thank the participants and staff of the Health Professionals Follow-up Study for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data.

Iowa-Mayo SPORE – NCI Specialized Programs of Research Excellence (SPORE) in Human Cancer (P50 CA97274); National Cancer Institute (P30 CA086862, P30 CA15083); Henry J. Predolin Foundation.

Italian GxE - Italian Association for Cancer Research (AIRC, Investigator Grant 11855) (PC); Fondazione Banco di Sardegna 2010–2012, and Regione Autonoma della Sardegna (LR7 CRP-S9812/2012) (MGE).

Mayo Clinic Case-Control – National Institutes of Health (R01 CA92153); National Cancer Institute (P30 CA015083).

MCCS – The Melbourne Collaborative Cohort Study recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 209057, 251553 and 504711 and by recurrent funding and infrastructure provided by Cancer Council Victoria. The incidence of malignancy and their participants’ vital status were ascertained through the Victorian Cancer Registry and the Australian Institute of Health and Welfare, including the National Death Index and the Australian Cancer Database.

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

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

NHS –The NHS was supported in part by National Institutes of Health grants CA186107, CA87969, CA49449, CA149445, and CA098122. We would like to thank the participants and staff of the Nurses’ Health Study for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data.

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

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

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

SCALE – Swedish Cancer Society (2009/659). Stockholm County Council (20110209) and the Strategic Research Program in Epidemiology at Karolinska Institutet. Swedish Cancer Society grant (02 6661). National Institutes of Health (5R01 CA69669-02); Plan Denmark.
UCSF2 – The UCSF studies were supported by the NCI, National Institutes of Health, CA1046282 and CA154643. The collection of cancer incidence data used in this study was supported by the California Department of Health Services as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute’s Surveillance, Epidemiology, and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention’s National Program of Cancer Registries, under agreement #1U58 DP000807-01 awarded to the Public Health Institute. The ideas and opinions expressed herein are those of the authors, and endorsement by the State of California, the California Department of Health Services, the National Cancer Institute, or the Centers for Disease Control and Prevention or their contractors and subcontractors is not intended nor should be inferred. 

UTAH - National Institutes of Health CA134674. Partial support for data collection at the Utah site was made possible by the Utah Population Database (UPDB) and the Utah Cancer Registry (UCR). Partial support for all datasets within the UPDB is provided by the Huntsman Cancer Institute (HCI) and the HCI Cancer Center Support grant, P30 CA42014. The UCR is supported in part by NIH contract HHSN261201000026C from the National Cancer Institute SEER Program with additional support from the Utah State Department of Health and the University of Utah.

YALE – National Cancer Institute (CA62006); National Cancer Institute (CA165923).

References

Genome wide association studies (GWAS) included in the evaluation of human leukocyte antigen (HLA) homozygosity and risk of four non-Hodgkin lymphoma (NHL) subtypes: diffuse large B-cell lymphoma (DLBCL), follicular lymphoma (FL), chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL), and marginal zone lymphoma (MZL).

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Study Abbreviation</th>
<th>GWAS Platform</th>
<th>NHL Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha-Tocopherol, Beta-Carotene Lung Cancer Prevention Study</td>
<td>ATBC</td>
<td>Illumina OmniExpress</td>
<td>43</td>
<td>17</td>
</tr>
<tr>
<td>British Columbia Non-Hodgkin Lymphoma Study</td>
<td>BCCA</td>
<td>Illumina OmniExpress</td>
<td>92</td>
<td>98</td>
</tr>
<tr>
<td>American Cancer Society Cancer Prevention Study-II Nutrition Cohort</td>
<td>CPS-II</td>
<td>Illumina OmniExpress</td>
<td>188</td>
<td>141</td>
</tr>
<tr>
<td>Treatment program of DLBCL patients from the Groupe d’Etude des Lymphomes de l’Adulte (GELA) consisting in LNH03-1B, 2B, 3B, 39B, 6B and 7B.</td>
<td>GELA</td>
<td>Illumina HumanHap 610K</td>
<td>549</td>
<td>0</td>
</tr>
<tr>
<td>Epidemiology &amp; Genetics Unit Lymphoma Case-Control study</td>
<td>ELCCS</td>
<td>Illumina OmniExpress</td>
<td>229</td>
<td>182</td>
</tr>
<tr>
<td>Environmental and genetic risks factors study in adult lymphoma</td>
<td>ENGELA</td>
<td>Illumina OmniExpress</td>
<td>56</td>
<td>30</td>
</tr>
<tr>
<td>European Prospective Investigation into Cancer, Chronic Diseases, Nutrition and Lifestyle</td>
<td>EPIC</td>
<td>Illumina OmniExpress</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Epilymph case-control study in six European countries</td>
<td>Epilymph</td>
<td>Illumina OmniExpress</td>
<td>198</td>
<td>123</td>
</tr>
<tr>
<td>Genetic Epidemiology of CLL (GEC) Consortium</td>
<td>GEC</td>
<td>Affymetrix 6.0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Health Professionals Follow-up Study</td>
<td>HPFS</td>
<td>Illumina OmniExpress</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Iowa-Mayo SPORE Molecular Epidemiology Resource</td>
<td>IOWA-MAYO SPORE</td>
<td>Illumina OmniExpress</td>
<td>146</td>
<td>228</td>
</tr>
<tr>
<td>Multicenter Italian study on gene-environment interactions in lymphoma etiology: translational aspects</td>
<td>Italian GxE</td>
<td>Illumina OmniExpress</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Study Name</td>
<td>Study Abbreviation</td>
<td>GWAS Platform</td>
<td>NHL Cases</td>
<td>Controls</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>--------------------</td>
<td>-----------------------------</td>
<td>--------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Mayo Clinic Case-Control Study of NHL and CLL</td>
<td>MAYO-Case-Control</td>
<td>Illumina OmniExpress</td>
<td>25</td>
<td>245</td>
</tr>
<tr>
<td>Mayo Clinic Case-Control Study of NHL and CLL</td>
<td>MAYO-Case-Control</td>
<td>Illumina HumanHap 660W</td>
<td>393</td>
<td>0</td>
</tr>
<tr>
<td>The Melbourne Collaborative Cohort Study</td>
<td>MCCC</td>
<td>Illumina OmniExpress</td>
<td>71</td>
<td>58</td>
</tr>
<tr>
<td>Memorial-Sloan Kettering Lymphoproliferative Disorders Study</td>
<td>MSKCC</td>
<td>Illumina OmniExpress</td>
<td>175</td>
<td>174</td>
</tr>
<tr>
<td>National Cancer Institute-Surveillance, Epidemiology, and End Results Interdisciplinary Case-Control Study of Non-Hodgkin’s Lymphoma</td>
<td>NCI-SEER</td>
<td>Illumina OmniExpress</td>
<td>251</td>
<td>217</td>
</tr>
<tr>
<td>Nurses’ Health Study</td>
<td>NHS</td>
<td>Illumina OmniExpress</td>
<td>28</td>
<td>24</td>
</tr>
<tr>
<td>New South Wales non-Hodgkin lymphoma study</td>
<td>NSW</td>
<td>Illumina OmniExpress</td>
<td>115</td>
<td>146</td>
</tr>
<tr>
<td>New York University Women’s Health Study</td>
<td>NYU-WHS</td>
<td>Illumina OmniExpress</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial</td>
<td>PLCO</td>
<td>Illumina OmniExpress</td>
<td>153</td>
<td>115</td>
</tr>
<tr>
<td>Scandinavian Lymphoma Epidemiology Study</td>
<td>SCALE</td>
<td>Illumina OmniExpress</td>
<td>405</td>
<td>0</td>
</tr>
<tr>
<td>Scandinavian Lymphoma Epidemiology Study</td>
<td>SCALE</td>
<td>Illumina HumanHap 317K</td>
<td>0</td>
<td>376</td>
</tr>
<tr>
<td>Molecular Epidemiology of non-Hodgkin lymphoma 1</td>
<td>UCSF1</td>
<td>Illumina OmniExpress</td>
<td>38</td>
<td>7</td>
</tr>
<tr>
<td>Molecular Epidemiology of non-Hodgkin lymphoma 1</td>
<td>UCSF1</td>
<td>Illumina HumanCNV370-Duo</td>
<td>254</td>
<td>210</td>
</tr>
<tr>
<td>Molecular Epidemiology of non-Hodgkin lymphoma 2</td>
<td>UCSF2</td>
<td>Illumina OmniExpress</td>
<td>0</td>
<td>119</td>
</tr>
<tr>
<td>Utah Chronic Lymphocytic Leukemia Study</td>
<td>UTAH</td>
<td>Illumina HumanHap 610K</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Population-based NHL case-control study in Connecticut women</td>
<td>YALE</td>
<td>Illumina OmniExpress</td>
<td>126</td>
<td>98</td>
</tr>
</tbody>
</table>
Table 2
Effect of homozygosity at the three HLA class I loci -A, -B and -C and three HLA class II loci -DRB1, DQB1, and DPB1 on susceptibility to four NHL subtypes (DLBCL, FL, CLL/SLL, and MZL) in Caucasian participants within participating lymphoma genome-wide association studies (analyses adjusted for sex, study or region, age, and ancestry/principal components).

<table>
<thead>
<tr>
<th>Class I locus</th>
<th>Controls (n=912)</th>
<th>DLBCL (n=3617)</th>
<th>FL (n=7888)</th>
<th>Controls (n=7441)</th>
<th>CLL/SLL (n=2078)</th>
<th>Controls (n=5991)</th>
<th>MZL (n=741)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-A Homozygote</td>
<td>6039</td>
<td>80</td>
<td>3096</td>
<td>86</td>
<td>1.00 (ref)</td>
<td>6843</td>
<td>88</td>
</tr>
<tr>
<td>HLA-B Homozygote</td>
<td>756</td>
<td>11</td>
<td>484</td>
<td>14</td>
<td>1.46 (0.98–1.46)</td>
<td>923</td>
<td>12</td>
</tr>
<tr>
<td>HLA-C Homozygote</td>
<td>643</td>
<td>93</td>
<td>3297</td>
<td>91</td>
<td>1.00</td>
<td>7330</td>
<td>93</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class II locus</th>
<th>HLA-DRB1</th>
<th>HLA-DQB1</th>
<th>HLA-DPB1</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-DRB1 Homozygote</td>
<td>6331</td>
<td>92</td>
<td>3173</td>
</tr>
<tr>
<td>HLA-DQB1 Homozygote</td>
<td>561</td>
<td>8</td>
<td>435</td>
</tr>
<tr>
<td>HLA-DPB1 Homozygote</td>
<td>6137</td>
<td>89</td>
<td>3055</td>
</tr>
</tbody>
</table>

| Total # of homozygous Class loci | 0 | 5535 | 80 | 2792 | 77 | 1.00 | 6266 | 80 | 2121 | 79 | 1.00 (ref) | 5965 | 80 | 2225 | 77 | 1.00 (ref) | 4805 | 80 | 586 | 79 | 1.00 (ref) |
| 1 | 950 | 14 | 524 | 14 | 1.48 (1.05–1.49) | 1120 | 14 | 361 | 64 | 0.98 (0.84–1.13) | 1009 | 14 | 457 | 70 | 1.19 (1.03–1.36) | 822 | 14 | 94 | 13 | 0.97 (0.76–1.23) |
| 2 | 297 | 4 | 187 | 5 | 1.30 (1.05–1.61) | 342 | 4 | 132 | 5 | 1.18 (0.93–1.51) | 323 | 4 | 130 | 5 | 1.08 (0.85–1.37) | 256 | 4 | 37 | 5 | 1.16 (0.80–1.68) |
| 3 | 130 | 2 | 114 | 3 | 1.30 (0.95–1.81) | 152 | 2 | 72 | 3 | 1.29 (0.93–1.79) | 144 | 2 | 66 | 2 | 1.16 (0.83–1.62) | 108 | 2 | 24 | 3 | 1.23 (0.95–1.58) |

<p>| p-trend | 0.6008 | 0.12 | 0.038 | 0.026 |
| OR per locus | 1.10 (0.83–1.19) | 1.06 (0.89–1.15) | 1.08 (1.00–1.16) | 1.10 (0.95–1.25) |</p>
<table>
<thead>
<tr>
<th></th>
<th>Controls (n=6912)</th>
<th>DLBCL (n=3617)</th>
<th>Controls (n=7880)</th>
<th>FL (n=2686)</th>
<th>Controls (n=7441)</th>
<th>CLL/SLL (n=2878)</th>
<th>Controls (n=5991)</th>
<th>MZL* (n=741)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n</strong></td>
<td>3972</td>
<td>1750</td>
<td>625</td>
<td>232</td>
<td>98</td>
<td>84</td>
<td>84</td>
<td>1</td>
</tr>
<tr>
<td><strong>%</strong></td>
<td>59</td>
<td>26</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td><strong>OR (95% CI)</strong></td>
<td>1.00 (ref)</td>
<td>1.11 (0.98–1.25)</td>
<td>1.11 (0.98–1.25)</td>
<td>1.11 (0.98–1.25)</td>
<td>1.11 (0.98–1.25)</td>
<td>1.11 (0.98–1.25)</td>
<td>1.11 (0.98–1.25)</td>
<td>1.11 (0.98–1.25)</td>
</tr>
<tr>
<td><strong>p-trend</strong></td>
<td>&lt;0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

*adjusted by geographic region (continent) rather than study

Cancer Res. Author manuscript available in PMC 2019 July 15.
Table 3

Effects of zygosity by individual HLA Class I and Class II loci, for DLBCL, MZL, FL, and CLL/SLL (analyses adjusted for sex, age, study/region, and ancestry/principal components).

<table>
<thead>
<tr>
<th>Class I locus</th>
<th>HLA-B</th>
<th>HLA-C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Controls (n=6912)</td>
<td>DLBCL (n=3687)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>HLA-C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homozygote</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterozygote</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class II locus</th>
<th>HLA-DRB1</th>
<th>HLA-DQB1</th>
<th>HLA-DPR1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Controls (n=6912)</td>
<td>DLBCL (n=3687)</td>
<td>Controls (n=7880)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>HLA-DPR1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homozygote</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterozygote</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*OR* and *p*-value were calculated in a logistic regression model with adjustment for sex, age, study/region, and ancestry/principal components.