**Infectious mononucleosis, immune genotypes, and non-Hodgkin lymphoma (NHL):**

**an InterLymph Consortium study.**

Niquelle Brown Wadé1, Cindy M. Chang2, David Conti1,3, Joshua Millstein1,3, Christine Skibola4, Alexandra Nieters5, Sophia Wang6, Silvia De Sanjose7,8, Eleanor Kane9, John Spinelli10, Paige Bracci11, Yawei Zhang12, Susan Slager13, Jun Wang1,2, Henrik Hjalgrim14, Karin Ekstrom Smedby15,Elizabeth E. Brown16, Ruth F. Jarrett17, Wendy Cozen1,3,18 **for the InterLymph Consortium Immunology and Infection Working Group.**

1 Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA

2 Division of Population Health Sciences, Center for Tobacco Products, Food and Drug Administration, Bethesda, Maryland, USA

3 USC Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, California, USA

4 Department of Hematology and Medical Oncology, Emory University School of Medicine, Atlanta, Georgia, USA

5 Center for Chronic Immunodeficiency (CCI), University Medical Center Freiburg, Germany, University of Freiburg, Freiburg, Germany

6 Department of Population Sciences, City of Hope National Medical Center, Duarte, California, USA

7 Sexual and Reproductive Health, PATH, Seattle, Washington, USA

8 CIBER Epidemiologia y Salud Publica, Barcelona, Spain

9 Department of Health Sciences, University of York, York, YO10 5DD, United Kingdom

10 Population Oncology, BC Cancer Agency; Epidemiology, Biostatistics and Public Health Practice, University of British Columbia, Vancouver, British Columbia, Canada

11Department of Epidemiology and Biostatistics, University of California at San Francisco, San Francisco, California, USA

12 Department of Surgery, Yale School of Medicine and Yale School of Public Health, New Haven, Connecticut, USA

13 Department of Epidemiology, Mayo Clinic, Rochester, Minnesota, USA

14 Department of Epidemiology Research, Statens Serum Institut, Copenhagen; Department of Hematology, Finsen Center, Rigshospitalet, Copenhagen, Denmark

15 Karolinska Instiutet, Karolinska University, Stockholm, Sweden University Hospital, Sweden

16 Department of Pathology and UAB Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, Alabama, USA

17 MRC-University of Glasgow Centre for Virus Research, Glasgow, Scotland, G61 1QH,

18 Department of Pathology, Keck School of Medicine, University of Southern California, Los Angeles California, USA

Disclaimer: The information in this article is not a formal dissemination of information by the FDA and does not represent agency position or policy. The contents are the responsibility of the authors alone.

Declaration of interests: This article was prepared while Dr. Cindy Chang was employed at the National Cancer Institute.

**Funding**

This study was supported in part by award number N01-CN-75014-20 and P30CA014089 from the National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

**Acknowledgements**

The authors would like to thank Annalie Landgren and Aaron Norman for supporting the InterLymph Consortium. In addition, the authors would like the acknowledge the support of the InterLymph Data Coordinating Center at Mayo Clinic.

**Abstract**

Background**:** We explored the interaction between non-Hodgkin lymphoma (NHL), infectious mononucleosis (IM) history, and immune-related genotypes in a pooled case-control analysis.
Methods**:** 7926 NHL patients and 10 018 controls from 12 studies were included. Self-reported IM history and genotypes were provided by the InterLymph Data Coordinating Center at Mayo Clinic. Odds ratios (OR) were estimated using multivariate logistic and linear regression, and interactions with the empirical Bayes method. Bonferroni corrections and pACT were used to account for multiple comparisons.
Results**:** There was evidence of an interaction effect between IM history and two variants on T-cell lymphoma (TCL) risk: rs1143627 in *interleukin-1B* (pinteraction= 0.02, ORinteraction = 0.09, 95% confidence interval [CI] = 0.01, 0.87) and rs1800797 in *interleukin-6* (pinteraction = 0.02, ORinteraction=0.08, 95% CI = 0.01, 0.80). Neither interaction effect withstood adjustment for multiple comparisons. Among controls, increasing socioeconomic status (OR = 1.69, 95% CI = 1.48, 1.93) and female sex (OR = 1.53, 95% CI = 1.26, 1.87) were positively associated with IM. Large sibship size (3+) was inversely associated with IM among controls born before 1960 (OR<1960 = 0.40, 95% CI = 0.24, 0.67), but not after.

Conclusions**:** Genetic risk variants in *IL1B* and *IL6* may affect the association between IM and TCL. Risk factors for IM are consistent with lower Epstein-Barr virus exposure in early life; the association with female sex is unexplained.

Keywords: Infectious mononucleosis, non-Hodgkin lymphoma, T-cell lymphoma, genotype, interleukin 1B, interleukin 6, siblings, family size, socioeconomic status

**Key Messages**

* A suggestion that genetic variation in *interleukin-1B (IL1B)* and *interleukin-6 (IL6)* may attenuate the association between infectious mononucleosis and T-cell lymphoma requires confirmation with larger numbers.
* Increasing socioeconomic status and female sex are independently associated with self-reported infectious mononucleosis.
* The number of siblings is inversely associated with self-reported infectious mononucleosis among those born before 1960 but not those born after.

## Introduction

Non-Hodgkin lymphoma (NHL) comprises a group of lymphoid malignancies with distinct histopathologies and risk patterns [1] originating from B- (~80%) and T-lymphocytes (~20%). Genetic or acquired immunodeficiency is the strongest risk factor, but more subtle immune alterations may also play a role in pathogenesis [2]. For example, there is a strong positive association between NHL and autoimmune disease [3,4] and an inverse association with atopy [5]. In addition to evidence of familiality for overall and subtype-specific NHL risk [6,7], variants in and near genes related to innate and adaptive immunity (*IL1RN, FCGR2A, TNFA*, HLA Class I and II) [8–10] have been implicated as potential risk factors.

Several infectious agents, including Epstein-Barr virus (EBV) [11], Hepatitis C virus [12], and Helicobacter pylori [13], contribute to NHL etiology through various mechanisms including direct transformation of lymphocytes, immunosuppression, chronic B-cell activation, and innate immune stimulation [14]. EBV, a ubiquitous member of the human herpesvirus family, induces B-cell growth by expression of viral proteins and non-coding RNAs [15]. The viral DNA persists as an episome in the host memory B-cell DNA after infection where it remains latent in the presence of a competent cytotoxic T-cell response. When acquired early in life, primary EBV infection is generally asymptomatic or causes a mild, non-specific, febrile illness [16]. In industrialized countries and populations of higher socioeconomic status (SES), primary infection is often delayed until adolescence or young adulthood. From 25% to 74% of those experiencing delayed primary infection develop infectious mononucleosis (IM), a moderate to severe clinical syndrome characterized by fever, tonsillar pharyngitis, and lymphadenopathy [17–19]. Propensity to develop the syndrome is influenced by genetic factors related to immune response [20,21].

In the largest pooled case-control study of NHL conducted to date from the International Lymphoma Epidemiology Consortium (InterLymph), Becker et al. observed a positive association between self-reported IM history and risk of all NHL (OR = 1.26, 95% CI [1.01, 1.57]). When stratified by subtype, associations were observed between IM and T-cell lymphoma (TCL) and a B-cell category combining chronic lymphocytic leukemia (CLL), small lymphocytic lymphoma (SLL), prolymphocytic lymphoma (PLL), and mantle cell lymphoma (MCL) [22].

Studies indicating familial IM clustering and higher concordance for IM risk among monozygotic compared to dizygotic twin pairs suggest a role for genetic susceptibility in IM etiology [23,24]. However, little is known about the influence of genetic factors on IM risk or their role in modifying the possible association between IM and NHL. Many of the genetic risk loci identified for NHL and NHL subtypes in previous InterLymph studies are in or near genes related to immune response that might also influence the association between IM and NHL risk [8,9,25–29].

In this InterLymph study, we examined the joint effects of IM history and 12 candidate immune-related variants on the risk of NHL, the impact of IM history on age at NHL diagnosis, and risk factors for IM among controls using previously collected demographic and familial information and the same candidate panel of immune-related genetic variants.

## Methods

### Study population

Participants included NHL patients and controls contributed from InterLymph Consortium member sites. All 12 studies (from 10 countries) had approval from their respective National or Institutional Review Boards, and participants provided signed informed consent according to the WMA Declaration of Helsinki Ethical Principles for Medical Research Involving Human Subjects in 1964. A summary of study details is provided inSupplemental Table 1 with additional details available in previous InterLymph publications [4,5,10,22,26–40].

InterLymph Consortium member studies were selected for inclusion based on the availability of self-reported IM history and candidate variant genotypes from at least 50% of participants. Participants who had missing data for age at enrollment, sex, SES, or IM history were excluded. Because the number of non-white participants in member studies was small and would require stratification for genetic analyses, we limited the study to white participants. Consistent with previous InterLymph analyses, participants who reported IM diagnosis less than 2 years before NHL diagnosis were excluded [22].

### Data collection

The InterLymph Data Coordinating Center (Mayo Clinic, Rochester, MN) harmonized data submitted by each study site into a de-identified, pooled dataset for analysis. Information on demographics, family structure (number of siblings and birth order), and IM history was self-reported using questionnaires [1]. Ethnicity/race was available for eleven of the twelve study centers included in the analysis, with the participants from most of these European, U.S., and Canadian studies being non-Hispanic white. Participants with missing race/ethnicity were included from SCALE (N=5683), Mayo Clinic (N=28), Yale (N=3), NCI-SEER-Seattle and Iowa (N=20) studies since the majority of the population in these study areas were non-Hispanic white; otherwise those with missing race were excluded. Socioeconomic status (SES) was categorized based on years of education (low: 0-12 years, high school or less; medium: 13-15 years, some college; high: 16+ years, college degree or more) or tertiles of the SES variable submitted by each individual study center.

The pooled analysis used existing genotype data on variants selected *a priori* based on results from previous functional analyses, association with NHL, or role in pro-/anti-inflammatory pathways [8,9,25–27]. The effects of these 12 genetic variants located in or near nine immune-response genes were assessed: *IL1A*-889C>T (rs1800587), *IL1B*–511C>T (rs16944), *IL1B*–31T>C (rs1143627), *IL1RN*–9589A>T (rs454078), *IL2*–384T>G (rs2069762), *IL6*–174G>C (rs1800795), *IL6*–597G>A (rs1800797), *IL10*–3575T>A (rs1800890), *IL10*–1082A>G (rs1800896), *TNF*–308G>A (rs1800629), *HLA class I* C>A (rs6457327), and *HLA class II* T>G (rs10484561). Genotyping was performed using either TaqMan (Applied Biosystems, Inc., Foster City, California), Pyrosequencing (Qiagen NV, Hilden, Germany), or Illumina Goldengate (Illumina, Inc., San Diego, California) genotyping assays. Additional technical details about genotyping methods used in each contributing study are included in previous publications [8,26,27,39,41].

All NHL diagnoses were confirmed by pathology report review, with the majority re-reviewed by a hematopathologist, depending on the study. NHL subtypes were classified according to the World Health Organization (WHO) classification in 2001 and 2008 [42–44] and include chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL: ICD-O-3 codes 9670, 9823), diffuse large B-cell lymphoma (DLBCL: 9679, 9680, 9684), follicular lymphoma (FL: 9690, 9691, 9695, 9698), mantle cell lymphoma (MCL: 9673), TCL (9702, 9705, 9708, 9709, 9714, 9716, 9717, 9718, 9719, 9729, 9827, 9834), and all NHL combined (defined by the above ICDO3 codes and 9671, 9675, 9687, 9689, 9699, 9700, 9701, 9728, 9826, 9832, 9833, 9591, and 9727). Patients with AIDS-related lymphomas were excluded.

### Statistical Analysis

*Candidate variants in linkage disequilibrium (LD):* SNP Annotation and Proxy Search (SNAP) [45] was used to assess LD via correlations between all pairs of candidate variants in the same gene.

*Main effect NHL associations:* Unconditional logistic regression was used to estimate odds ratios (ORs) and 95% confidence intervals (CIs) for the association between IM and NHL and for associations between candidate genetic variants and NHL. Consistent with other InterLymph publications [26,27], all genetic variants were coded as dichotomous variables assuming a dominant model (absence or presence of minor allele). All models were adjusted for age at NHL diagnosis/enrollment, sex, study center, and SES.

*Gene-environment interaction in NHL risk:* The effect of interaction between IM and immune-related genotypes on NHL risk was assessed using the empirical Bayes approach described by Mukherjee et al. [46]. Sensitivity analyses were then performed using unconditional logistic regression to test the association between IM and NHL stratified by each candidate variant genotype. Likelihood ratio tests were used to estimate p-values for interactions against a baseline model assuming multiplicativity of odds ratios. Models were adjusted for the covariates listed above. Associations were examined for all NHL combined and by NHL subtype.

*Association with age at NHL diagnosis:* We performed a case-only analysis stratified by NHL subtype to assess the association between IM and age at NHL diagnosis using linear regression models adjusted for sex, study center, SES, and year of birth (to account for a potential birth cohort effect).

*Risk factors for IM among controls:* We evaluated the associations between IM and demographic factors (SES and sex), family structure (birth order and sibship size), and genetic factors (candidate genetic variants mentioned above) using unconditional multivariable logistic regression models adjusted for age and study center. Models including SES, family structure, or genetic variants as the exposure variable of interest were adjusted for sex. Models including sex or genetic factors as the exposure variable of interest were further adjusted for SES. Analyses with family structure as the exposure variable of interest were stratified by year of birth (1901-1960, 1961-1990 based on a historical shift in attitudes and policies about preschool attendance in the 1960s [47]) and restricted to sites with year of birth data for controls (Mayo Clinic; British Columbia, Canada; and all EpiLymph sites). All analyses of risk factors for IM were restricted to controls.

*Goodness of fit, sensitivity analysis, and multiple comparisons:* All genetic data were assessed for deviations from allele frequencies expected under Hardy-Weinberg equilibrium among controls, and a sensitivity analysis was conducted in which we excluded study centers from the analysis of the specific genetic variants for which within-center allele frequencies were inconsistent with Hardy-Weinberg equilibrium at p<0.05. Hosmer-Lemeshow goodness-of-fit tests were performed to assess all logistic regression models [48], and residual analysis was conducted to ensure linear regression models satisfied the appropriate assumptions. Additional sensitivity analyses were conducted excluding studies using clinic-based control recruitment methods.

All statistical tests were two-sided. For genetic analyses, the pACT statistic was used to account for multiple comparisons and correlated tests from variants within the same region [49]. For analyses of demographic and familial factors, a Bonferroni correction was used to account for multiple comparisons. Uncorrected p-values are reported in tables. For those associations with uncorrected p-values <0.05, Bonferroni-corrected p-values (pBon) or pACT statistics are noted in the text. Statistical analysis was performed using Stata, version 13 (StataCorp, LP, College Station, TX).

## Results

### Main NHL associations

7926 NHL patients and 10 018 controls from 12 InterLymph studies met the inclusion criteria. The distribution of NHL patients and controls by selected demographic and clinical characteristics is shown in Table 1. The majority (83%) of patients were diagnosed with mature B-cell lymphoma (Table 2); the remainder were diagnosed with mature T-cell, precursor cell, and not otherwise specified (NOS) lymphomas.

### Analysis with SNAP indicated candidate risk variants in IL1B (r2IL1B: rs16944, rs1143627 = 0.96) and in IL6 (r2IL6: rs1800795, rs1800797 = 0.97) were in high LD, and candidate variants in IL10 were in moderate LD (r2IL10: rs1800890, rs1800896 = 0.66).

After adjustment for multiple comparisons, we observed strong main effects for associations between *HLA* variants and NHL (pACT<0.001 and pACT=0.004), an *IL1RN* variant and NHL (pACT=0.04), IM and CLL/SLL (pBon=0.04), and IM and MCL (pBon=0.01) (Supplemental Table 2). A history of IM was associated with all NHL combined, CLL/SLL, and MCL (Supplemental Table 3). The direction of the association between IM and NHL risk was consistent when restricted to population-based studies (not shown in tables). Thus, the main effects of genotype and IM for associations with all NHL and NHL subtypes were largely consistent with previously reported results from a subset of the same InterLymph studies [8,9,22,25–28].

### Gene-environment interaction in NHL risk

There was an interaction effect between a genetic variant in the *IL1B* gene (rs1143627T) and IM history on TCL (ORinteraction=0.09, 95% CI=0.01-0.87, p=0.02), DLBCL (ORinteraction=0.61, 95% CI=0.34-1.08, p=0.04), and all NHL combined (ORinteraction=0.76, 95% CI=0.57-1.02, p=0.03) risk. An interaction between rs1800797A in the *IL6* gene and TCL (ORinteraction =0.08, 95% CI= 0.01-0.80, p=0.02) (Table 3). None of the associations persisted after adjustment for multiple comparisons (pACT>0.05). These results were directionally consistent when restricted to population-based studies.

### Age at NHL diagnosis

Although self-reported history of IM was strongly associated with age at NHL diagnosis among NHL patients of each subtype, the association was not present after adjusting for birth year, suggesting a cohort effect (Supplemental Table 4). We observed similar results when the analysis was restricted to population-based studies.

### IM risk factors among controls

Among 10 018 control participants for whom IM history was available, 521 reported a positive history of IM. Increasing SES level (ORtrend = 1.69, 95% CI = 1.48, 1.93, pBon<0.001) and female sex (OR = 1.53, 95% CI = 1.26, 1.87, pBon<0.001) were positively associated with IM (Table 4). These results were consistent when restricted to participants recruited using population-based methods. An association between sibship size and IM history among all controls did not withstand adjustment for multiple comparisons (Table 5, OR3+ siblings = 0.57, 95% CI = 0.38, 0.85, pBon>0.05). However, stratification of controls by year of birth revealed a strong inverse association between large sibship size (3+ siblings) and IM among those born before 1960 (pBon<0.001). There was no evidence of an association between birth order and IM risk in controls with 2 or more siblings (Table 5). None of the candidate variants showed evidence of an association with IM history (Table 6).

## Discussion

Infectious mononucleosis was associated with an increased risk of TCL in the original main effects InterLymph paper [22] and with a 32% (p = 0.17) increased risk among our subset of InterLymph participants. The minor allele in variant rs1143627 in the promoter region of the *IL1B* gene appeared to attenuate the effect of IM on TCL and DLBCL risk (with a much stronger magnitude for TCL), although the effects for both subtypes did not persist after adjustment for multiple comparisons. A similar interaction effect was observed for all NHL combined and is likely attributable to the preponderance of the DLBCL subtype among our sample of NHL patients.

 IL-1B, the cytokine encoded by this gene, is an inflammatory response and fever mediator, and contributes to several lymphocyte activities including growth and differentiation of B-cells [50], proliferation of T-helper Type 2 (Th2) clones [51], and activation of Th17 cells [52]. We observed a suggestive association of similar magnitude between rs16944, an *IL1B* variant highly correlated with rs1143627, and TCL. IL1B is required for T-cell activation in some immune responses [53,54] and thus could contribute to increased T-cell replication. The minor alleles of the two variants examined in our study are associated with lower expression of IL1B [55] and may decrease T-cell activation in the setting of increased EBV load. This decrease in activation may, in turn, attenuate the effect of IM. rs16944 has also been associated with uncontrolled EBV replication in liver transplant patients, who later develop post-transplant lymphoproliferative disorder [56], suggesting a link between IL-1B and dysfunctional control of EBV. There was also suggestive association between the functional variant rs1800797 in the *IL6* gene promoter region and risk of TCL. Through complex interactions with nearby variants, rs1800797 regulates the gene that encodes the inflammatory cytokine IL6, which influences growth and differentiation of T-cells, among many other immune functions [57,58]. Follow-up of these observations in a targeted study is warranted because of the potential biological pathway.

Among controls, increasing SES was associated with elevated risk of IM; the risk of IM was roughly two times higher among high SES participants compared to low SES participants. This observation is consistent with previous studies, which have suggested high SES is a surrogate for a lower probability of EBV exposure in early life (due to fewer siblings and less crowded environments) and thus, a higher risk of developing IM [18,59–61]. In our study, we observed a strong relationship between large sibship size and IM (Table 5, pBon=0.001) among controls born before 1960 but not in those born after 1960. Controls born after 1960 may have been more likely to attend preschool, which would provide EBV exposure in early life and simulate a large family. Alternatively, it may be that after 1960, overcrowding decreased, and hygienic behaviors generally increased, mitigating the importance of family size. The findings for our controls born before 1960 are consistent with previous reports of inverse associations between sibship size and IM [62,63]. There was no evidence of an association between birth order and IM history among controls with 2 or more siblings; however, the sample size of controls born after 1960 with larger sibship size was insufficient to calculate stable estimates of effect size (Table 5) because a number of study sites did not collect year of birth data for controls.

In our study, IM was also more common among females (6% prevalence) than males (5% prevalence), in both cases and controls. In a prospective study among university students, Crawford et al. did not find a difference in IM prevalence among EBV-seroconverters by gender [64]; however higher rates of hospitalization for IM among teen and young adult females in the UK were reported by ﻿Ramagopalan et al. [65]. While Crawford et al. measured propensity for IM upon primary EBV infection (i.e. seroconverters) among college-age subjects, our study relied on self-reported IM history among all subjects regardless of age, including IM that developed prior to and after college. In addition, the denominator in the Crawford study included only EBV-seronegative college students, while our denominator included adults of all ages, some of whom surely acquired EBV prior to adolescence and were therefore not at risk of IM. Thus, our results are not necessarily discrepant with those from the Crawford study. Explanations for the higher IM prevalence among females in our study include lower rates of EBV infection among females in childhood producing a higher pool at risk for IM in adolescence and young adulthood, a higher reporting of symptomatic disease [66–68], higher likelihood of medical care seeking behavior and thus diagnosis, or a true biological effect enhancing IM risk in females compared to males at a wider range of ages (young or older than college). Sex-based biological differences in response to infections have been reported. For example, females mount more vigorous antibody- and cell-mediated immune responses following some infections and vaccines than men [69,70].

A limitation of our study is reliance on self-reported IM history, which could be affected by recall bias. However, IM is a severe and debilitating syndrome of relatively long duration, interrupting young adult life; therefore, it is unlikely that a participant would forget this experience. Another limitation is the use of cross-sectional data for determining IM risk factors which prevents the establishment of a temporal relationship between IM and some demographic variables (e.g. SES) and relies on the assumption that SES measured at the time of the study reflects SES in adolescence and young adulthood when IM would have occurred. These types of information biases are likely to have resulted in non-differential misclassification of the exposure with respect to the outcome, biasing results toward the null. In addition, the strongly positive trends in the expected direction (SES, sibship size) suggest a true association.

Although the results can be generalized to adults of European descent living in the United States and Europe, the limited number of ethnically diverse participants enrolled in these studies and the exclusion of HIV/AIDS-related lymphomas and post-transplant lymphomas limits generalizability to other groups. Because NHL patients were recruited after the onset of disease, those with longer post-diagnosis survival times were more likely to enroll in the study and complete questionnaires. This ascertainment bias prevents us from generalizing to NHL patients with very short survival times, although rapid case ascertainment methods at individual study sites dampened the impact of this bias. Furthermore, data from sites using clinic-based recruitment methods for enrolling controls are subject to Berkson’s bias since patient controls are likely to be sicker than the general population from which cases were ascertained. Many admitting conditions of clinic-based controls may have some immune component which can obscure the effect of immune-related genetic variants on NHL and IM. Results of sensitivity analyses excluding clinic-based sites were directionally consistent with results using the full dataset, indicating the effect of Berkson’s bias on our study results was minimal.

Our study was underpowered to detect an interaction between uncommon variants and IM within rare NHL subtype strata after adjusting for multiple comparisons. For example, in order to achieve 80% power of detecting an interaction odds ratio of 0.09 for rs1143627 at α = 0.05 after accounting for multiple comparisons, we would have needed 518 genotyped TCL patients. Thus, even with the overall large numbers of cases and controls in the study, there was inadequate power to detect associations by subtype.

In summary, we confirmed the long-standing association between IM risk and higher SES and lower sibship size (for those born prior to 1960) and showed a female excess that requires confirmation and further explanation. This study was also the first to explore possible interaction between immune response genotypes and IM history on NHL risk [71]. Although we observed a possible interaction that affected the risk of a rare NHL subtype, our study was underpowered to overcome multiple comparisons. Confirmation will require a well-characterized, targeted study with larger numbers.

## Tables

Table 1: Demographic characteristics of NHL patients and controls

Table 2: Subtypes among NHL patients

Table 3: Interaction between IM history and candidate risk variants [*IL1A* (rs1800587), *IL1B* (rs16944, rs1143627), *IL1RN* (rs454078), *IL2* (rs2069762), *IL6* (rs1800795, rs1800797), *IL10* (rs1800896, rs1800890), *TNFA* (rs1800629), *HLA I* (rs6457327), and *HLA II* (rs10484561)] by NHL subtype: empirical-Bayes estimates of interaction effects

Table 4: Associations between IM history and demographic/familial factors among controls

Table 5: Association between IM history and family structure among controls by year of birth

Table 6: Associations between IM and candidate risk variants [*IL1A* (rs1800587), *IL1B* (rs16944, rs1143627), *IL1RN* (rs454078), *IL2* (rs2069762), *IL6* (rs1800795, rs1800797), *IL10* (rs1800896, rs1800890), *TNFA* (rs1800629), *HLA I* (rs6457327), and *HLA II* (rs10484561)] among controls

Supplemental Table 1: Source of subjects

Supplemental Table 2: Associations between NHL and candidate risk variants [*IL1A* (rs1800587), *IL1B* (rs16944, rs1143627), *IL1RN* (rs454078), *IL2* (rs2069762), *IL6* (rs1800795, rs1800797), *IL10* (rs1800896, rs1800890), *TNFA* (rs1800629), *HLA I* (rs6457327), and *HLA II* (rs10484561)]

Supplemental Table 3: Association between NHL and infectious mononucleosis

Supplemental Table 4: Association between age at NHL diagnosis and infectious mononucleosis

Supplemental Table 5: Interaction between IM history and candidate risk variants [*IL1A* (rs1800587), *IL1B* (rs16944, rs1143627), *IL1RN* (rs454078), *IL2* (rs2069762), *IL6* (rs1800795, rs1800797), *IL10* (rs1800896, rs1800890), *TNFA* (rs1800629), *HLA I* (rs6457327), and *HLA II* (rs10484561)]

|  |
| --- |
| **Table 1: Demographic characteristics of NHL patients and controls** |
|   |  |  Controls (N=10018)  |   | NHL Patients (N=7926) |
|  |  | Negative IM history |  | Positive IM History |  | Negative IM history |  | Positive IM History |
|  |  | N | (%) |  | N | (%) |  | N | (%) |  | N | (%) |
| Study Center | *BC* | 604 | (6%) |  | 35 | (7%) |  | 566 | (8%) |  | 42 | (10%) |
| *EpiLymph-Czech Republic* | 289 | (3%) |  | 8 | (2%) |  | 165 | (2%) |  | 5 | (1%) |
| *EpiLymph-France* | 250 | (3%) |  | 5 | (1%) |  | 198 | (3%) |  | 3 | (1%) |
| *EpiLymph-Germany* | 628 | (7%) |  | 21 | (4%) |  | 435 | (6%) |  | 18 | (4%) |
| *EpiLymph-Ireland* | 198 | (2%) |  | 5 | (1%) |  | 116 | (2%) |  | 11 | (3%) |
| *EpiLymph-Italy* | 331 | (3%) |  | 3 | (1%) |  | 177 | (2%) |  | 2 | (0%) |
| *EpiLymph-Spain* | 603 | (6%) |  | 5 | (1%) |  | 418 | (6%) |  | 6 | (1%) |
| *Mayo Clinic* | 1,014 | (11%) |  | 85 | (16%) |  | 779 | (10%) |  | 80 | (20%) |
| *NCI-SEER* | 378 | (4%) |  | 25 | (5%) |  | 543 | (7%) |  | 48 | (12%) |
| *Scale* | 2,830 | (30%) |  | 106 | (20%) |  | 2,653 | (35%) |  | 94 | (23%) |
| *UCSF* | 1,752 | (18%) |  | 189 | (36%) |  | 946 | (13%) |  | 63 | (15%) |
| *Yale* | 620 | (7%) |  | 34 | (7%) |  | 523 | (7%) |  | 35 | (9%) |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| SES | *Low* | 3,282 | (35%) |  | 63 | (12%) |  | 3,037 | (40%) |  | 74 | (18%) |
|  | *Medium* | 3,169 | (33%) |  | 146 | (28%) |  | 2,400 | (32%) |  | 134 | (33%) |
|  | *High* | 3,046 | (32%) |  | 312 | (60%) |  | 2,082 | (28%) |  | 199 | (49%) |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Birth order | *First/Only* | 3,318 | (35%) |  | 182 | (35%) |  | 2,555 | (34%) |  | 149 | (37%) |
|  | *2nd* | 2,412 | (25%) |  | 154 | (30%) |  | 1,781 | (24%) |  | 115 | (28%) |
|  | *3rd* | 1,278 | (13%) |  | 83 | (16%) |  | 1,031 | (14%) |  | 50 | (12%) |
|  | *4th* | 1,653 | (17%) |  | 57 | (11%) |  | 1,392 | (19%) |  | 40 | (10%) |
|  | *Missing* | 836 | (9%) |  | 45 | (9%) |  | 760 | (10%) |  | 53 | (13%) |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Number of Siblings | *0* | 394 | (4%) |  | 24 | (5%) |  | 255 | (3%) |  | 23 | (6%) |
| *1* | 1,578 | (17%) |  | 110 | (21%) |  | 1,144 | (15%) |  | 71 | (17%) |
| *2* | 2,147 | (23%) |  | 159 | (31%) |  | 1,603 | (21%) |  | 118 | (29%) |
| *3* | 4,673 | (49%) |  | 189 | (36%) |  | 3,908 | (52%) |  | 153 | (38%) |
| *Missing* | 705 | (7%) |  | 39 | (7%) |  | 609 | (8%) |  | 42 | (10%) |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Sex | *Male* | 5,018 | (53%) |  | 260 | (50%) |  | 4,052 | (54%) |  | 186 | (46%) |
|  | *Female* | 4,479 | (47%) |  | 261 | (50%) |  | 3,467 | (46%) |  | 221 | (54%) |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | Mean ± SD | Med (IQR) |  | Mean ± SD | Med (IQR) |  | Mean ± SD | Med (IQR) |  | Mean ± SD | Med (IQR) |
| Age at NHL Diagnosis/Interview | 57 ± 15 | 60 (21) |   | 46 ± 15 | 47 (22) |   | 60 ± 12 | 62 (17) |   | 52 ± 13 | 53 (19) |
| IM: infectious mononucleosis |
| SES: socioeconomic status |
| NHL: non-Hodgkin lymphoma |
| SD: standard deviation |
| IQR: interquartile range |

|  |
| --- |
| **Table 2: Subtypes among NHL Patients** |
|   |  | N | (%) |
| B-cell | *DLBCL* | 2246 |  (28%) |
|  | *CLL/SLL/PLL/MCL* | 1466 |  (18%) |
|  | *Follicular* | 1691 |  (21%) |
|  | *MZL* | 447 |  (6%) |
|  | *MCL* | 325 |  (4%) |
|  | *LPL/Waldenstrom* | 228 |  (3%) |
|  | *Hairy cell* | 75 |  (1%) |
|  | *Burkitt* | 63 |  (1%) |
|  | *Precursor B-cell* | 40 |  (1%) |
|  | *Burkitt-like* | 27 |  (0.3%) |
|  | *PLL* | 4 |  (0.05%) |
|  | *B-Cell NOS* | 534 |  (7%) |
|  | *TOTAL B-Cell* | 7146 |  (90%) |
|  |  |  |  |
| T-Cell | *Peripheral T-cell* | 262 |  (3%) |
|  | *MF/SS* | 166 |  (2%) |
|  | *Precursor T-cell* | 26 |  (0.3%) |
|  | *Nasal NK* | 17 |  (0.2%) |
|  | *Large granular* | 7 |  (0.1%) |
|  | *T-PLL* | 4 |  (0.1%) |
|  | *T-Cell NOS* | 27 |  (0.3%) |
|  | *TOTAL T-Cell* | 509 |  (6%) |
|  |  |  |  |
| NOS |  | 210 |  (3%) |
| Missing |  | 61 |  (1%) |

|  |
| --- |
| **Table 3: Interaction between IM history and candidate risk variants [IL1A (rs1800587), IL1B (rs1143627), IL1RN (rs454078), IL2 (rs2069762), IL6 (rs1800795, rs1800797), IL10 (rs1800890), TNFA (rs1800629), HLA I (rs6457327), and HLA II (rs10484561)] by NHL subtype: empirical-Bayes estimates of interaction effects** |
| NHL Subtype |  | Genotyped Controls | Genotyped NHL Patients | Interaction OR a | Interaction 95% CI |  |
| Variant | N | N | p-value b |
| All NHL | IL1A-889C>T (rs1800587) | 2,317 | 2,084 | 1.13 | [0.69, 1.87] | 0.69 |
|  | IL1B-511C>T (rs16944) | 1,280 | 1,311 | 0.66 | [0.35, 1.22] | 0.09 |
|  | IL1B-31T>C (rs1143627) | 3,715 | 3,130 | **0.76** | **[0.57, 1.02]** | **0.03** |
|  | IL1RN-9589A>T (rs454078) | 2,319 | 2,068 | 1.02 | [0.62, 1.67] | 0.53 |
|  | IL2–384T>G (rs2069762) | 2,320 | 2,080 | 0.99 | [0.62, 1.57] | 0.48 |
|  | IL6-174G>C (rs1800795) | 2,347 | 2,099 | 1.08 | [0.77, 1.52] | 0.68 |
|  | IL6-597G>A (rs1800797) | 3,852 | 3,304 | 1.10 | [0.81, 1.49] | 0.72 |
|  | IL10-1082A>G (rs1800896) | 4,173 | 3,472 | 0.81 | [0.59, 1.10] | 0.09 |
|  | IL10-3575T>A (rs1800890) | 5,914 | 5,629 | 0.80 | [0.57, 1.13] | 0.10 |
|  | TNF-308G>A (rs1800629) | 5,562 | 5,546 | 0.86 | [0.59, 1.27] | 0.23 |
|  | HLA: C>A (rs6457327) | 2,963 | 2,457 | 1.07 | [0.65, 1.76] | 0.61 |
|  | HLA: T>G (rs10484561) | 3,989 | 3,176 | 0.93 | [0.62, 1.40] | 0.37 |
|  |  |  |  |  |  |  |
| CLL/SLL | IL1A-889C>T (rs1800587) | 2,317 | 366 | 1.03 | [0.45, 2.35] | 0.53 |
|  | IL1B-511C>T (rs16944) | 1,280 | 117 | 1.30 | [0.22, 7.56] | 0.62 |
|  | IL1B-31T>C (rs1143627) | 3,715 | 646 | 0.96 | [0.49, 1.87] | 0.45 |
|  | IL1RN-9589A>T (rs454078) | 2,319 | 364 | 1.81 | [0.80, 4.07] | 0.92 |
|  | IL2–384T>G (rs2069762) | 2,320 | 365 | 1.77 | [0.79, 3.98] | 0.92 |
|  | IL6-174G>C (rs1800795) | 2,347 | 364 | 1.13 | [0.49, 2.57] | 0.61 |
|  | IL6-597G>A (rs1800797) | 3,852 | 666 | 0.89 | [0.45, 1.76] | 0.37 |
|  | IL10-1082A>G (rs1800896) | 4,173 | 669 | 0.80 | [0.39, 1.62] | 0.27 |
|  | IL10-3575T>A (rs1800890) | 5,914 | 1,204 | 0.68 | [0.37, 1.24] | 0.10 |
|  | TNF-308G>A (rs1800629) | 5,562 | 1,186 | 0.89 | [0.47, 1.71] | 0.37 |
|  | HLA: C>A (rs6457327) | 2,963 | 389 | 1.05 | [0.36, 3.02] | 0.54 |
|  | HLA: T>G (rs10484561) | 3,989 | 623 | 0.54 | [0.20, 1.50] | 0.12 |
|  |  |  |  |  |  |  |
| DLBCL | IL1A-889C>T (rs1800587) | 2,317 | 541 | 0.98 | [0.48, 2.01] | 0.48 |
|  | IL1B-511C>T (rs16944) | 1,280 | 384 | 0.75 | [0.34, 1.68] | 0.24 |
|  | IL1B-31T>C (rs1143627) | 3,715 | 877 | **0.61** | **[0.34, 1.08]** | **0.04** |
|  | IL1RN-9589A>T (rs454078) | 2,319 | 530 | 0.83 | [0.41, 1.68] | 0.30 |
|  | IL2–384T>G (rs2069762) | 2,320 | 538 | 2.02 | [0.99, 4.13] | 0.97 |
|  | IL6-174G>C (rs1800795) | 2,347 | 537 | 0.83 | [0.43, 1.60] | 0.29 |
|  | IL6-597G>A (rs1800797) | 3,852 | 922 | 0.92 | [0.52, 1.64] | 0.39 |
|  | IL10-1082A>G (rs1800896) | 4,173 | 928 | 1.10 | [0.58, 2.09] | 0.61 |
|  | IL10-3575T>A (rs1800890) | 5,914 | 1,496 | 0.74 | [0.45, 1.21] | 0.11 |
|  | TNF-308G>A (rs1800629) | 5,562 | 1,447 | 0.72 | [0.42, 1.23] | 0.11 |
|  | HLA: C>A (rs6457327) | 2,963 | 701 | 0.66 | [0.32, 1.37] | 0.13 |
|  | HLA: T>G (rs10484561) | 3,989 | 840 | 1.05 | [0.51, 2.17] | 0.55 |
|  |  |  |  |  |  |  |
| FL | IL1A-889C>T (rs1800587) | 2,317 | 527 | 1.18 | [0.58, 2.41] | 0.68 |
|  | IL1B-511C>T (rs16944) | 1,280 | 331 | 0.53 | [0.20, 1.36] | 0.09 |
|  | IL1B-31T>C (rs1143627) | 3,715 | 706 | 0.98 | [0.54, 1.77] | 0.47 |
|  | IL1RN-9589A>T (rs454078) | 2,319 | 526 | 0.63 | [0.31, 1.29] | 0.10 |
|  | IL2–384T>G (rs2069762) | 2,320 | 528 | 0.69 | [0.35, 1.35] | 0.14 |
|  | IL6-174G>C (rs1800795) | 2,347 | 533 | 1.34 | [0.69, 2.60] | 0.80 |
|  | IL6-597G>A (rs1800797) | 3,852 | 757 | 1.78 | [0.94, 3.39] | 0.96 |
|  | IL10-1082A>G (rs1800896) | 4,173 | 750 | 0.67 | [0.37, 1.20] | 0.09 |
|  | IL10-3575T>A (rs1800890) | 5,914 | 1,125 | 0.89 | [0.51, 1.54] | 0.34 |
|  | TNF-308G>A (rs1800629) | 5,562 | 1,130 | 0.91 | [0.50, 1.65] | 0.38 |
|  | HLA: C>A (rs6457327) | 2,963 | 510 | 1.65 | [0.72, 3.79] | 0.88 |
|  | HLA: T>G (rs10484561) | 3,989 | 696 | 0.86 | [0.46, 1.62] | 0.32 |
|  |  |  |  |  |  |  |
| MCL | IL1A-889C>T (rs1800587) | 2,317 | 103 | 2.24 | [0.48, 10.33] | 0.85 |
|  | IL1B-511C>T (rs16944) | 1,280 | 61 | 0.25 | [0.02, 3.07] | 0.14 |
|  | IL1B-31T>C (rs1143627) | 3,715 | 146 | 1.35 | [0.33, 5.57] | 0.66 |
|  | IL1RN-9589A>T (rs454078) | 2,319 | 102 | 1.24 | [0.28, 5.53] | 0.61 |
|  | IL2–384T>G (rs2069762) | 2,320 | 103 | 0.74 | [0.16, 3.43] | 0.35 |
|  | IL6-174G>C (rs1800795) | 2,347 | 105 | 0.28 | [0.06, 1.24] | 0.05 |
|  | IL6-597G>A (rs1800797) | 3,852 | 159 | 0.29 | [0.08, 1.11] | 0.04 |
|  | IL10-1082A>G (rs1800896) | 4,173 | 171 | 0.70 | [0.15, 3.22] | 0.32 |
|  | IL10-3575T>A (rs1800890) | 5,914 | 285 | 0.59 | [0.20, 1.77] | 0.17 |
|  | TNF-308G>A (rs1800629) | 5,562 | 279 | 3.27 | [1.06, 10.05] | 0.98 |
|  | HLA: C>A (rs6457327) | 2,963 | 116 | 2.75 | [0.34, 22.05] | 0.83 |
|  | HLA: T>G (rs10484561) | 3,989 | 158 | 0.29 | [0.04, 2.33] | 0.12 |
|  |  |  |  |  |  |  |
| T-Cell | IL1A-889C>T (rs1800587) | 2,317 | 127 | 2.76 | [0.36, 21.42] | 0.83 |
|  | IL1B-511C>T (rs16944) | 1,280 | 97 | 0.01 | [0.00, 2.55] | 0.05 |
|  | IL1B-31T>C (rs1143627) | 3,715 | 206 | **0.09** | **[0.01, 0.87]** | **0.02** |
|  | IL1RN-9589A>T (rs454078) | 2,319 | 125 | 0.37 | [0.05, 2.75] | 0.16 |
|  | IL2–384T>G (rs2069762) | 2,320 | 127 | 0.71 | [0.12, 4.36] | 0.36 |
|  | IL6-174G>C (rs1800795) | 2,347 | 129 | 0.05 | [0.00, 1.01] | 0.03 |
|  | IL6-597G>A (rs1800797) | 3,852 | 219 | **0.08** | **[0.01, 0.80]** | **0.02** |
|  | IL10-1082A>G (rs1800896) | 4,173 | 233 | 0.89 | [0.16, 5.04] | 0.45 |
|  | IL10-3575T>A (rs1800890) | 5,914 | 378 | 1.07 | [0.36, 3.16] | 0.55 |
|  | TNF-308G>A (rs1800629) | 5,562 | 369 | 1.20 | [0.39, 3.71] | 0.62 |
|  | HLA: C>A (rs6457327) | 2,963 | 183 | 3.36 | [0.48, 23.47] | 0.89 |
|   | HLA: T>G (rs10484561) | 3,989 | 210 | 1.82 | [0.31, 10.58] | 0.75 |
| NHL: non-Hodgkin Lymphoma. CLL/SLL: chronic lymphocytic leukemia/small lymphocytic lymphoma. DLBCL: diffuse large B-cell lymphoma. FL: follicular lymphoma. MCL: mantle cell lymphoma. |
| OR: odds ratio. |
| CI: confidence interval. |
| a Interaction ORs, CIs, and p-values calculated using empirical-Bayes method adjusted for age, sex, study center, and socioeconomic status. |
| b Significant values are shown in bold but did not retain significance after accounting for multiple comparisons using pACT statistic.  |

|  |
| --- |
| **Table 4: Associations between IM history and demographic/familial factors among 10 018 controls** |
|  |  | IM negative controls | IM positive controls |  |  |  |  |
|  |  | N | N | OR a | 95% CI | p-value |  |
| SES | Low SES | 3,282 | 63 | (ref) | -- |  |  |
|  | Med SES | 3,169 | 146 | **1.69** | **[1.24, 2.30]** | **0.001** | b |
|  | High SES | 3,046 | 312 | **2.86** | **[2.13, 3.82]** | **<.001** | b |
|  | Trend | 9,497 | 521 | **1.69** | **[1.48, 1.93]** | **<.001** | b |
|  |  |  |  |  |  |  |  |
| Sex | Male | 5,018 | 260 | (ref) | -- |  |  |
|   | Female | 4,479 | 261 | **1.53** | **[1.26, 1.87]** | **<.001** | b |
| IM: infectious mononucleosis |
| SES: socioeconomic status |
| OR: odds ratio. |
| CI: confidence interval. |
| FDR: false discovery rate. |
| a ORs, CIs, and p-values calculated using logistic regression models adjusted for age and study center. Models for SES, number of siblings, and birth order were further adjusted for sex. Model for sex was further adjusted for SES. Results with p<0.05 are shown in bold.  |
| b Strong evidence against the null hypotheses (no association) after accounting for multiple comparisons using Bonferroni correction. |

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 5: Association between IM history and family structure among controls with available year of birth data** |  |  |  |
|  |  |  | IM negative controls | IM positive controls |  |  |  |  |  |  |
| Year of birth a |   | N | N | OR d | 95% CI | p-value |  |  |  |
| All | Number of Siblings b | 0 or 1 | 933 | 54 | (ref) | -- |  |  |  |  |
|  |  | 2 | 899 | 56 | 1.03 | [0.69, 1.55] | 0.88 |  |  |  |
|  |  | 3+ | 2,067 | 56 | **0.57** | **[0.38, 0.85]** | **0.01** |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  | Birth Order c | 1st | 750 | 30 | (ref) | -- |  |  |  |  |
|  | (among controls with 2+ siblings) | 2nd | 745 | 36 | 1.41 | [0.85, 2.36] | 0.19 |  |  |  |
|  | 3rd+ | 1,447 | 46 | 0.98 | [0.60, 1.59] | 0.92 |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 1901 - 1960 | Number of Siblings b | 0 or 1 | 733 | 34 | (ref) | -- |  |  |  |  |
|  |  | 2 | 685 | 34 | 1.10 | [0.66, 1.82] | 0.71 |  |  |  |
|  |  | 3+ | 1,822 | 35 | **0.40** | **[0.24, 0.67]** | **<.001** | e |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  | Birth Order c | 1st | 602 | 21 | (ref) | -- |  |  |  |  |
|  | (among controls with 2+ siblings) | 2nd | 640 | 22 | 1.04 | [0.56, 1.94] | 0.89 |  |  |  |
|  | 3rd+ | 1,244 | 26 | 0.61 | [0.34, 1.12] | 0.11 |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 1961-1990 | Number of Siblings b | 0 or 1 | 200 | 20 | (ref) | -- |  |  |  |  |
|  |  | 2 | 213 | 22 | 1.01 | [0.52, 1.97] | 0.99 |  |  |  |
|  |  | 3+ | 245 | 21 | 1.07 | [0.53, 2.17] | 0.85 |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  | Birth Order c | 1st | 147 | 9 | (ref) | -- |  |  |  |  |
|  | (among controls with 2+ siblings) | 2nd | 105 | 14 | -- | -- |  |  |  |  |
|   | 3rd+ | 203 | 20 | -- | -- |   |   |  |  |
| IM: infectious mononucleosis |  |  |
| SES: socioeconomic status |  |  |
| OR: odds ratio. |  |  |
| CI: confidence interval. |  |  |
| a Year of birth was reported for controls was reported at eight study sites (38% of controls) |  |  |
| b Sibship size was reported for 99.6% of controls at sites with year of birth data |  |  |
| c Birth order was not reported for 1% of controls at sites with year of birth data |  |  |
| d ORs, CIs, and p-values calculated using logistic regression models adjusted for age, study center, and sex. Results with p<0.05 are shown in bold.  |  |  |
| e Strong evidence against the null hypotheses (no association) after accounting for multiple comparisons using Bonferroni correction. |  |  |

|  |
| --- |
| **Table 6: Associations between IM and candidate risk variants [IL1A (rs1800587), IL1B (rs1143627), IL1RN (rs454078), IL2 (rs2069762), IL6 (rs1800797), IL10 (rs1800890), TNFA (rs1800629), HLA I (rs6457327), and HLA II (rs10484561)] among controls with available genetic data** |
|  | All studies |
|  | IM Negative Controls | IM Positive Controls |  |  |  |
|  | N | N | OR a | 95% CI | p-value b |
| IL1A-889C>T (rs1800587) | 2,168 | 149 | 0.72 | [0.51, 1.02] | 0.07 |
| IL1B-511C>T (rs16944) | 1,213 | 67 | 0.76 | [0.46, 1.27] | 0.30 |
| IL1B-31T>C (rs1143627) | 3,533 | 182 | 1.03 | [0.75, 1.40] | 0.86 |
| IL1RN-9589A>T (rs454078) | 2,169 | 150 | 0.76 | [0.53, 1.08] | 0.13 |
| IL2–384T>G (rs2069762) | 2,173 | 147 | 1.25 | [0.88, 1.76] | 0.22 |
| IL6-174G>C (rs1800795) | 2,195 | 152 | 0.94 | [0.66, 1.34] | 0.73 |
| IL6-597G>A (rs1800797) | 3,662 | 190 | 0.95 | [0.69, 1.30] | 0.73 |
| IL10-1082A>G (rs1800896) | 3,975 | 198 | 1.04 | [0.75, 1.46] | 0.81 |
| IL10-3575T>A (rs1800890) | 5,663 | 251 | 1.16 | [0.88, 1.53] | 0.29 |
| TNF-308G>A (rs1800629) | 5,318 | 244 | 1.22 | [0.92, 1.62] | 0.16 |
| HLA: C>A (rs6457327) | 2,858 | 105 | 1.22 | [0.80, 1.86] | 0.36 |
| HLA: T>G (rs10484561) | 3,801 | 188 | 0.88 | [0.59, 1.31] | 0.52 |
| a Adjusted for age, sex, study center, socioeconomic status |
| b No evidence against the null hypotheses (no association). |

## WORKS CITED

1. Morton LM, Sampson JN, Cerhan JR, et al. Rationale and Design of the International Lymphoma Epidemiology Consortium (InterLymph) Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst - Monogr* 2014;**2014**:1–14.

2. Grulich AE, Vajdic CM, Cozen W. Altered Immunity as a Risk Factor for Non-Hodgkin Lymphoma. *Cancer Epidemiol Prev Biomarkers* 2007;**16**:405–9.

3. Morton LM, Wang SS, Cozen W, et al. Etiologic heterogeneity among non-Hodgkin lymphoma subtypes. *Blood* 2008;**112**:5150–60.

4. Smedby KE, Vajdic CM, Falster M, et al. Autoimmune disorders and risk of non-Hodgkin lymphoma subtypes: A pooled analysis within the InterLymph Consortium. *Blood* 2008;**111**:4029–38.

5. Cozen W, Cerhan JR, Martinez-Maza O, et al. The effect of atopy, childhood crowding, and other immune-related factors on non-Hodgkin lymphoma risk. *Cancer Causes Control* 2007;**18**:821–31.

6. Goldin LR, Björkholm M, Kristinsson SY, et al. Highly increased familial risks for specific lymphoma subtypes. *Br J Haematol* 2009;**146**:91–4.

7. Wang SS, Flowers CR, Kadin ME, et al. Medical history, lifestyle, family history, and occupational risk factors for peripheral T-cell lymphomas: The interlymph non-hodgkin lymphoma subtypes project. *J Natl Cancer Inst - Monogr* 2014;**2014**:66–75.

8. Hosgood HD, Purdue MP, Wang SS, et al. A pooled analysis of three studies evaluating genetic variation in innate immunity genes and non-Hodgkin lymphoma risk. *Br J Haematol* 2011;**152**:721–6.

9. Skibola CF, Akers NK, Conde L, et al. Multi-locus HLA class I and II allele and haplotype associations with follicular lymphoma. *Tissue Antigens* 2012;**79**:279–86.

10. Wang SS, Abdou AM, Morton LM, et al. Human leukocyte antigen class I and II alleles in non-Hodgkin lymphoma etiology. *Blood* 2010;**115**:4820–3.

11. Thorley-Lawson DA, Gross A. Persistence of the Epstein–Barr Virus and the Origins of Associated Lymphomas. *N Engl J Med* 2004;**350**:1328–37.

12. Ferri C, Caracciolo F, Zignego AL, et al. Hepatitis C virus infection in patients with non-Hodgkin’s lymphoma. *Br J Haematol* 1994;**88**:392–4.

13. Parsonnet J, Hansen S, Rodriguez L, et al. Helicobacter pylori infection and gastric lymphoma. *N Engl J Med* 1994;**330**:1267–71.

14. Engels E a. Infectious agents as causes of non-Hodgkin lymphoma. *Cancer Epidemiol Biomarkers Prev* 2007;**16**:401–4.

15. Longnecker RM, Kieff E, Cohen JI. Epstein-Barr Virus. In: Knipe DM, Howley PM, editors. Fields Virol. 6th ed., Philadelphia, PA: Lippincott Williams & Wilkins; 2013, p. 1898–959.

16. Fleisher G, Henle W, Henle G, et al. Primary Infection with Epstein-Barr Virus in Infants in the United States: Clinical and Serologic Observations. *J Infect Dis* 1979;**139**:553–8.

17. Luzuriaga K, Sullivan JL. Infectious mononucleosis. *N Engl J Med* 2010;**362**:1993–2000.

18. Macsween KF, Johannessen I. Epstein-Barr Virus (EBV): Infectious Mononucleosis and Other Non-malignant EBV-Associated Diseases. In: Kaslow RA, Stanberry LR, Le Duc JW, editors. Viral Infect Humans Epidemiol Control, Boston, MA: Springer US; 2014, p. 867–96.

19. Balfour HH, Dunmire SK, Hogquist KA. Infectious mononucleosis. *Clin Transl Immunol* 2015;**4**:e33.

20. Tian C, Hromatka BS, Kiefer AK, et al. Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. *Nat Commun* 2017;**8**:599.

21. Mcaulay K a, Higgins CD, Macsween KF, et al. HLA class I polymorphisms are associated with development of infectious mononucleosis upon primary EBV infection. *J Clin Invest* 2007;**117**.

22. Becker N, Falster MO, Vajdic CM, et al. Self-reported history of infections and the risk of non-Hodgkin lymphoma: An InterLymph pooled analysis. *Int J Cancer* 2012;**131**:2342–8.

23. Hwang AE, Hamilton AS, Cockburn MG, et al. Evidence of genetic susceptibility to infectious mononucleosis: a twin study. *Epidemiol Infect* 2012;**140**:2089–95.

24. Rostgaard K, Wohlfahrt J, Hjalgrim H. A genetic basis for infectious mononucleosis: evidence from a family study of hospitalized cases in Denmark. *Clin Infect Dis* 2014;**58**:1684–9.

25. Kane E, Skibola CF, Bracci PM, et al. Non-Hodgkin Lymphoma, Body Mass Index, and Cytokine Polymorphisms: A Pooled Analysis from the InterLymph Consortium. *Cancer Epidemiol Biomarkers Prev* 2015;**24**:1061–70.

26. Skibola CF, Bracci PM, Nieters A, et al. Tumor necrosis factor (TNF) and lymphotoxin-alpha (LTA) polymorphisms and risk of non-Hodgkin lymphoma in the InterLymph Consortium. *Am J Epidemiol* 2010;**171**:267–76.

27. Rothman N, Skibola CF, Wang SS, et al. Genetic variation in TNF and IL10 and risk of non-Hodgkin lymphoma: A report from the InterLymph Consortium. *Lancet Oncol* 2006;**7**:27–38.

28. Cerhan JR, Ansell SM, Fredericksen ZS, et al. Genetic variation in 1253 immune and inflammation genes and risk of non-Hodgkin lymphoma. *Blood* 2007;**110**:4455–63.

29. Cerhan JR, Fredericksen ZS, Novak AJ, et al. 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 Epidemiol Biomarkers Prev* 2012;**21**:1799–806.

30. Besson H, Brennan P, Becker N, et al. Tobacco smoking, alcohol drinking and non-Hodgkin’s lymphoma: A European multicenter case-control study (Epilymph). *Int J Cancer* 2006;**119**:901–8.

31. Becker N, Fortuny J, Alvaro T, et al. Medical history and risk of lymphoma: Results of a European case-control study (EPILYMPH). *J Cancer Res Clin Oncol* 2009;**135**:1099–107.

32. Cerhan JR, Fredericksen ZS, Wang AH, et al. Design and validity of a clinic-based case-control study on the molecular epidemiology of lymphoma. *Int J Mol Epidemiol Genet* 2011;**2**:95–113.

33. Chatterjee N, Hartge P, Cerhan JR, et al. Risk of non-Hodgkin’s lymphoma and family history of lymphatic, hematologic, and other cancers. *Cancer Epidemiol Biomarkers Prev* 2004;**13**:1415–21.

34. Hughes AM, Armstrong BK, Vajdic CM, et al. Sun exposure may protect against non-Hodgkin lymphoma: a case-control study. *Int J Cancer* 2004;**112**:865–71.

35. Smedby KE, Hjalgrim H, Melbye M, et al. Ultraviolet radiation exposure and risk of malignant lymphomas. *J Natl Cancer Inst* 2005;**97**:199–209.

36. Holly EA, Lele C, Bracci PM, et al. Case-control study of non-Hodgkin’s lymphoma among women and heterosexual men in the San Francisco Bay Area, California. *Am J Epidemiol* 1999;**150**:375–89.

37. Holly EA, Bracci PM. Population-based study of non-Hodgkin lymphoma, histology, and medical history among human immunodeficiency virus-negative participants in San Francisco. *Am J Epidemiol* 2003;**158**:316–27.

38. Morton LM, Holford TR, Leaderer B, et al. Alcohol use and risk of non-Hodgkin’s lymphoma among Connecticut women (United States). *Cancer Causes Control* 2003;**14**:687–94.

39. Conde L, Halperin E, Akers NK, et al. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. *Nat Genet* 2010;**42**:661–4.

40. Spinelli JJ, Ng CH, Weber J-P, et al. Organochlorines and risk of non-Hodgkin lymphoma. *Int J Cancer* 2007;**121**:2767–75.

41. Smedby KE, Foo JN, Skibola CF, et al. GWAS of follicular lymphoma reveals allelic heterogeneity at 6p21.32 and suggests shared genetic susceptibility with diffuse large B-cell lymphoma. *PLoS Genet* 2011;**7**.

42. Morton LM, Turner JJ, Cerhan JR, et al. Proposed classification of lymphoid neoplasms for epidemiologic research from the Pathology Working Group of the International Lymphoma Epidemiology Consortium (InterLymph). *Blood* 2007;**110**:695–708.

43. Turner JJ, Morton LM, Linet MS, et al. InterLymph hierarchical classification of lymphoid neoplasms for epidemiologic research based on the WHO classification (2008): Update and future directions. *Blood* 2010;**116**.

44. Swerdlow SH, Campo E, Harris NL, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. vol. 2. 4th ed. International Agency for Research on Cancer; 2008.

45. Johnson AD, Handsaker RE, Pulit SL, et al. SNAP: A web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008;**24**:2938–9.

46. Mukherjee B, Ahn J, Gruber SB, et al. Tests for gene-environment interaction from case-control data: a novel study of type I error, power and designs. *Genet Epidemiol* 2008;**32**:615–26.

47. Kamerman SB. A global history of early childhood education and care. 2006.

48. Lemeshow S, Hosmer DW. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982;**115**:92–106.

49. Conneely KN, Boehnke M. So Many Correlated Tests, So Little Time! Rapid Adjustment of P Values for Multiple Correlated Tests. *Am J Hum Genet* 2007;**81**:1158–68.

50. Pike BL, Nossal GJ. Interleukin 1 can act as a B-cell growth and differentiation factor. *Proc Natl Acad Sci U S A* 1985;**82**:8153–7.

51. Lichtman AH, Chin J, Schmidt JA, et al. Role of interleukin 1 in the activation of T lymphocytes. *Proc Natl Acad Sci U S A* 1988;**85**:9699–703.

52. Schett G, Dayer J-M, Manger B. Interleukin-1 function and role in rheumatic disease. *Nat Rev Rheumatol* 2016;**12**:14–24.

53. Nambu A, Nakae S, Iwakura Y. IL-1β, but not IL-1α, is required for antigen-specific T cell activation and the induction of local inflammation in the delayed-type hypersensitivity responses. *Int Immunol* 2006;**18**:701–12.

54. Hackett RJ, Davis LS, Lipsky PE. Comparative effects of tumor necrosis factor-alpha and IL-1 beta on mitogen-induced T cell activation. *J Immunol* 1988;**140**:2639–44.

55. Hirbod-Mobarakeh A, Amirzargar AA, Nikbin B, et al. Immunogenetics of Cancer. Cancer Immunol, Berlin, Heidelberg: Springer Berlin Heidelberg; 2015, p. 295–341.

56. Kasztelewicz B, Jankowska I, Pawłowska J, et al. The impact of cytokine gene polymorphisms on Epstein-Barr virus infection outcome in pediatric liver transplant recipients. *J Clin Virol* 2012;**55**:226–32.

57. Terry CF, Loukaci V, Green FR. Cooperative influence of genetic polymorphisms on interleukin 6 transcriptional regulation. *J Biol Chem* 2000;**275**:18138–44.

58. Burger R. Impact of Interleukin-6 in hematological malignancies. *Transfus Med Hemotherapy* 2013;**40**:336–43.

59. Nye FJ. Social class and infectious mononucleosis. *J Hyg (Lond)* 1973;**71**:145–9.

60. Lehane DE. A Seroepidemiologic Study of Infectious Mononucleosis. *JAMA* 1970;**212**:2240.

61. Sumaya C V, Henle W, Henle G, et al. Seroepidemiologic study of Epstein-Barr virus infections in a rural community. *J Infect Dis* 1975;**131**:403–8.

62. Rostgaard K, Nielsen TR, Wohlfahrt J, et al. Sibship structure and risk of infectious mononucleosis: a population-based cohort study. *Int J Epidemiol* 2014;**43**:1607–14.

63. Liu Z, Fang F, Chang ET, et al. Sibship size, birth order and risk of nasopharyngeal carcinoma and infectious mononucleosis: a nationwide study in Sweden. *Int J Epidemiol* 2016;**45**:825–34.

64. Crawford DH, Macsween KF, Higgins CD, et al. A cohort study among university students: identification of risk factors for Epstein-Barr virus seroconversion and infectious mononucleosis. *Clin Infect Dis* 2006;**43**:276–82.

65. Ramagopalan S V, Giovannoni G, Yeates DG, et al. Sex ratio of infectious mononucleosis and possible relevance to multiple sclerosis. *Mult Scler J* 2013;**19**:359–61.

66. Petersen I, Thomas JM, Hamilton WT, et al. Risk and predictors of fatigue after infectious mononucleosis in a large primary-care cohort. *QJM - Mon J Assoc Physicians* 2006;**99**:49–55.

67. Buchwald DS, Rea TD, Katon WJ, et al. Acute infectious mononucleosis: characteristics of patients who report failure to recover. *Am J Med* 2000;**109**:531–7.

68. Macsween KF, Higgins CD, McAulay K a, et al. Infectious mononucleosis in university students in the United Kingdom: evaluation of the clinical features and consequences of the disease. *Clin Infect Dis* 2010;**50**:699–706.

69. Beagley KW, Gockel CM. Regulation of innate and adaptive immunity by the female sex hormones oestradiol and progesterone. *FEMS Immunol Med Microbiol* 2003;**38**:13–22.

70. Fink AL, Klein SL. Sex and Gender Impact Immune Responses to Vaccines Among the Elderly. *Physiology* 2015;**30**:408–16.

71. Houldcroft CJ, Kellam P. Host genetics of Epstein-Barr virus infection, latency and disease. *Rev Med Virol* 2015;**25**:71–84.