A pooled investigation of Toll-like receptor gene variants and risk of non-Hodgkin lymphoma

Mark P. Purdue, Qing Lan, Sophia S. Wang, Anne Kricker, Idan Menashe, Tong-Zhang Zheng, Patricia Hartge, Andrew E. Grulich, Yawei Zhang, Lindsay M. Morton, Claire M. Vajdic, Theodore R. Holford, Richard K. Severson, Brian P. Leaderer, James R. Cerhan, Meredith Yeager, Wendy Cozen, Kevin Jacobs, Scott Davis, Nathaniel Rothman, Stephen J. Chanock, Nilanjana Chatterjee and Bruce K. Armstrong

Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Rockville, MD 20892, USA, School of Public Health, University of Sydney, Sydney 2006, Australia, Department of Epidemiology and Public Health, Yale School of Medicine, New Haven, CT 06520, USA, National Centre in HIV Epidemiology and Clinical Research, University of New South Wales, Sydney 2010, Australia, University of New South Wales Cancer Research Centre, Prince of Wales Clinical School, University of New South Wales, Sydney 2052, Australia, Department of Family Medicine and Karmanos Cancer Institute, Wayne State University, Detroit, MI 48201, USA, Mayo Clinic, College of Medicine, Rochester, MN 55905, USA, Core Genotyping Facility, Advanced Technology Center, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Gaithersburg, MD 20877, USA, Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA 90033, USA and Fred Hutchinson Cancer Research Center and University of Washington, Seattle, WA 98109, USA

Notes: To whom correspondence should be addressed. Tel: +1 301 451 5036; Fax: +1 301 402 1819; Email: purdue@lab.nih.gov

Toll-like receptors (TLRs) may influence the development of non-Hodgkin lymphoma (NHL) given their important roles in recognizing microbial pathogens and stimulating multiple immune pathways. We conducted an investigation of TLR gene variants in a pooled analysis including three population-based case-control studies of NHL (1946 cases and 1808 controls). Thirty-six tag single-nucleotide polymorphisms (SNPs) in TLR2, TLR4 and the TLR10–TLR1–TLR6 gene cluster were genotyped. Two TLR10–TLR1–TLR6 variants in moderate linkage disequilibrium were significantly associated with NHL: rs10008492 [odds ratio for CT genotype (OR_CT) 1.12, 95% confidence interval (CI) 0.97–1.30; OR TT 1.40, 95% CI 1.15–1.71; \( P_{\text{trend}} = 0.001 \)] and rs4833103 (OR_R 0.75, 95% CI 0.64–0.88; OR_A 0.74, 95% CI 0.62–0.90; \( P_{\text{trend}} = 0.002 \); \( P_{\text{dominant}} = 0.0002 \)). Associations with these SNPs were consistent across all the three studies and did not appreciably differ by histologic subtype. We found little evidence of association between TLR2 variation and all NHL, although the rare variant rs3804100 was significantly associated with marginal zone lymphoma (MZL). Both overall (OR_TCDC 1.89, 95% CI 1.27–2.81; \( P_{\text{dominant}} = 0.002 \)) and in two of the three studies. No associations with TLR4 variants were observed. This pooled analysis provides strong evidence that variation in the TLR10–TLR1–TLR6 region is associated with NHL risk and suggests that TLR2 variants may influence susceptibility to MZL.

Introduction

The etiology of non-Hodgkin lymphoma (NHL) is poorly understood, although various lines of evidence suggest that infections and immune dysregulation play a role. This evidence includes elevated incidence rates in immunosuppressed people and among individuals with selected autoimmune diseases, established links between specific infectious agents and rare NHL subtypes [Epstein-Barr virus and Burkitt lymphoma, human T-cell leukemia/lymphoma virus 1 and adult T-cell leukemia/lymphoma, human herpesvirus 8 and primary effusion lymphoma, Helicobacter pylori and gastric mucosa-associated lymphoid tissue (MALT) lymphoma] and associations with NHL risk for allergies and indicators of early-life exposure to infectious agents (1). In light of these findings, it is plausible that dysregulation in the recognition of and immune response to microbial pathogens may influence the risk of NHL development.

Toll-like receptors (TLRs) are a family of transmembrane receptors that play a key role in mounting an immune response against microbial pathogens. These receptors have been evolutionarily conserved to recognize specific microbial molecular components; once activated, TLRs engage a signaling cascade resulting in the stimulation of innate and adaptive immune responses targeting the invading pathogen. To date, 10 functional human TLRs have been identified (2). TLR-2 and TLR-4 are the principal receptors involved in the recognition of bacterial cell wall components; TLR-4 binds to Gram-negative bacterial lipopolysaccharides, whereas TLR-2 recognizes other cell wall molecules from a wide variety of bacteria as well as capsid ligands from several viruses (2,3). The capability of TLR-2 to recognize such a broad spectrum of ligands is the result of its ability to form heterodimer receptor complexes with TLR-1 and -6. Some endogenous molecules may also trigger TLR-mediated inflammatory responses; suspected TLR ligands include heat shock proteins, HMGB1, fibrogen and host DNA (4).

TLRs have been hypothesized as possible contributors to lymphoma and other immunologic/inflammatory diseases given their important role in recognizing microbial ligands and stimulating multiple immune pathways (2,5–6). Single-nucleotide polymorphisms (SNPs) in TLR2 (rs16933783; \( r^2 = 0.46 \)) and TLR4 (Arg509Gly; rs4986790) have been associated with lymphoma risk in recent studies (5–7). TLR1 and TLR6, located along with TLR10 [a TLR of unknown function (8)] in a 57 kb long region on 4p14, have not been studied in detail in relation to NHL, although an uncommon TLR6 variant was significantly associated with increased risk of NHL and selected subtypes in one study (9). Variation elsewhere within the TLR10–TLR1–TLR6 gene cluster has been associated with other diseases, including asthma (10,11) and prostate cancer (12).

To further explore whether genetic variation in TLR genes influences NHL pathogenesis, we conducted a pooled investigation of tag SNPs in TLR2, TLR4 and the TLR10–TLR1–TLR6 region in three population-based case–control studies of NHL.

Materials and methods

Study population

Our study population was derived from pooling three independent population-based case–control studies, which have been described in detail previously: the National Cancer Institute (NCI)-Surveillance Epidemiology and End Results (SEER) NHL case–control study (13,14), the Connecticut NHL case–control study (15,16) and the New South Wales (NSW) NHL case–control study (17,18). Selected characteristics for each study are presented in Table I. All the three studies included first primary NHL cases only, and population controls were frequency matched to cases. The protocols for each study were approved by the following institutions: the Institutional Review Boards of the NCI and each SEER center for the NCI-SEER study; Yale University, the Connecticut Department of Public Health and the NCI for the Connecticut.

Abbreviations:
- CI: confidence interval
- CLL: chronic lymphocytic leukemia
- DLBCL: diffuse large B-cell lymphoma
- FL: follicular lymphoma
- LD: linkage disequilibrium
- MALT: mucosa-associated lymphoid tissue
- MZL: marginal zone lymphoma
- NCI: National Cancer Institute
- NHL: non-Hodgkin lymphoma
- NSW: New South Wales
- OR: odds ratio
- SEER: Surveillance Epidemiology and End Results
- SLL: small lymphocytic lymphoma
- SNP: single-nucleotide polymorphism
- TLR: Toll-like receptor

© The Author 2008. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org
study and all participating institutions for the NSW study. All study participants provided informed consent.

**NHL pathology classification**
All cases were histologically confirmed by the local diagnosing pathologist in the NCI-SEER study and by central review by two independent pathologists in the Connecticut study. In the NSW study, all cases were histologically confirmed by the local diagnosing pathologist and a confirmatory central pathology review. Our studies primarily included SLL rather than CLL cases because these disorders were not considered the same entity until the World Health Organization classification was introduced in 2001 (20). A lymphoma/multiple myeloma case was considered a lymphoma if the last exon was binned using a binning threshold of $t^{2} > 0.80$. When there were multiple transcripts available for genes, the primary transcript was assessed.

**Biological samples and DNA extraction**
Study participants who did not provide a biologic specimen, did not have sufficient material for DNA extraction or sufficient DNA for genotyping or whose genotyped sex was discordant from the questionnaire data were excluded from this analysis. As described in Materials and Methods, the final analytic population consisted of 1946 cases and 1808 controls (NCI-SEER, 990/828; Connecticut, 956/980; and SEER area 99/98). For the NCI-SEER study, DNA was extracted from buffy coats using QIAamp DNA Blood Midi Kits by laboratory staff at the Viral Epidemiology Section, SAIC-Frederick, NCI-Frederick. TLR10, TLR2, and TLR4 were assessed in a subset of the NCI-SEER study. Genotyping was performed at the National Cancer Institute Core Genotyping Laboratory (v1.1: http://tagzilla.nci.nih.gov/), which employs the pairwise binning algorithm of Carlson et al. (24). For each gene, SNPs within the region 20 kbp of the ATG-translation initiation codon and 10 kbp of the end of the last exon were binned using a binning threshold of $t^{2} > 0.80$. When there were multiple transcripts available for genes, the primary transcript was assessed. Genotyping was performed at the National Cancer Institute Core Genotyping Facility (Gaithersburg, MD) using the Illumina GoldenGate platform, with the 36 tag SNPs included in an oligo pool assay (OPA) investigating genetic variants from different candidate pathways. Data for an additional non-synonymous TLR4 SNP (D299G; rs4986790) that had been previously genotyped by TaqMan assay in the NCI-SEER (14) and Connecticut studies were also included in the pooled analysis, increasing the total number of investigated SNPs to 37. Gene coverage (the number of genotyped SNPs divided by the number of SNPs from the design set of SNPs in Hamap Build 20 CEU samples) was calculated to be 61, 69 and 94% for TLR10–TLR5, TLR6, TLR2 and TLR4, respectively. All investigated SNPs in TLR1, TLR2, TLR4, TLR6 and TLR10 had genotype concordance rates and assay completion rates >99%. For each SNP, an exact test of Hardy–Weinberg proportions was performed among Caucasian controls. One TLR2 SNP deviated significantly from Hardy–Weinberg proportions ($t^{2} = 4666187, P = 0.02$).

We excluded samples with OPA-wide sample completion rates <90% (NCI-SEER: 11 cases, 6 controls; Connecticut: 2 controls; NSW: 4 cases, 9 controls). The final analytic population from pooling the three studies consisted of 1946 cases and 1808 controls (NCI-SEER, 990/828; Connecticut, 436/515; NSW, 520/465).

**Statistical methods**
Statistical analyses were performed using SAS version 9.1 (SAS Institute, Cary, NC) unless otherwise specified. Odds ratios (ORs) and 95% confidence intervals (CIs) estimating the relative risk of NHL in relation to SNP genotype were calculated using unconditional logistic regression adjusting for age, race/ethnicity, sex and study center. Analyses restricted to non-Hispanic Caucasians were also performed. Tests for trend (i.e. assuming a codominant genetic model) were performed by modeling the number of rare alleles (0, 1 or 2) as

---

**Table I. Selected characteristics of the NCI-SEER, Connecticut and NSW NHL case–control studies**

<table>
<thead>
<tr>
<th>Location</th>
<th>NCI-SEER</th>
<th>Connecticut</th>
<th>NSW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>Residents of the Iowa, Detroit, Los Angeles and Seattle SEER registries</td>
<td>Residents of the Connecticut SEER registry</td>
<td>Residents of the Australian State of NSW or the Australian Capital Territory via the NSW Central Cancer Registry</td>
</tr>
<tr>
<td>Age range (years)</td>
<td>20–74</td>
<td>21–84</td>
<td>20–74</td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>Excluded known HIV-positive individuals</td>
<td>Excluded males</td>
<td>Excluded known HIV-positive individuals and organ transplant recipients</td>
</tr>
<tr>
<td>Histopathologic classification system</td>
<td>ICD-O-2</td>
<td>REAL</td>
<td>WHO + ICD-O-3</td>
</tr>
<tr>
<td>Control selection</td>
<td>&lt;65 years: random digit dialing</td>
<td>&lt;65 years: random digit dialing</td>
<td>Electoral rolls</td>
</tr>
<tr>
<td>Matching variables</td>
<td>Age (5 year groups), sex, race and SEER area</td>
<td>Age (5 year groups)</td>
<td>Age (5 year groups), sex, state or territory</td>
</tr>
<tr>
<td>Study population (participation)</td>
<td>Cases: $n = 1521$ (76%) Controls: $n = 1057$ (52%)</td>
<td>Cases: $n = 601$ (72%) Controls: $n = 717$ (&lt;65 years: 69%; ≥65 years: 47%)</td>
<td>Cases: $n = 694$ (85%) Controls: $n = 694$ (61%)</td>
</tr>
<tr>
<td>Risk factor information</td>
<td>Self-administered questionnaire, in-person interview</td>
<td>Self-administered questionnaire, in-person interview</td>
<td>Self-administered questionnaire, telephone interview</td>
</tr>
<tr>
<td>DNA source</td>
<td>Venous blood or mouthwash buccal cell sample</td>
<td>Venous blood or mouthwash buccal cell sample</td>
<td>Venous blood sample</td>
</tr>
<tr>
<td>Restrictions for this analysis</td>
<td>Excluded participants who provided a buccal cell sample</td>
<td>Excluded participants who were not considered the same entity until the World Health Organization classification was introduced in 2001</td>
<td></td>
</tr>
<tr>
<td>Genotyped for this analysis</td>
<td>Cases: $n = 1001$ Controls: $n = 834$</td>
<td>Cases: $n = 436$ Controls: $n = 517$</td>
<td>Cases: $n = 524$ Controls: $n = 474$</td>
</tr>
</tbody>
</table>

---

**HIV, human immunodeficiency virus; ICD-O-2/3, International Classification of Diseases for Oncology, second/third revision; REAL, Revised European American Lymphoma system; WHO, World Health Organization.**

*a* Participation was defined as the percentage interviewed among those approached. *b* Participation was defined as the percentage approached among those approached. **Selected characteristics of the NCI-SEER, Connecticut and NSW NHL case–control studies**

<table>
<thead>
<tr>
<th>SNP</th>
<th>TLR10</th>
<th>TLR2</th>
<th>TLR4</th>
</tr>
</thead>
</table>
| Allele frequency | $0.05$ genotyped in the Caucasian (CEU) population sample of the Hamap Project (Data Release 20/Phase II, NCBI B35 assembly, dbSNP b125). Tag SNP selection was done using the software application TagZilla (v1.1: http://tagzilla.nci.nih.gov/), which employs the pairwise binning algorithm of Carlson et al. (24). For each gene, SNPs within the region 20 kbp of the ATG-translation initiation codon and 10 kbp of the end of the last exon were binned using a binning threshold of $t^{2} > 0.80$. When there were multiple transcripts available for genes, the primary transcript was assessed. Genotyping was performed at the National Cancer Institute Core Genotyping Facility (Gaithersburg, MD) using the Illumina GoldenGate platform, with the 36 tag SNPs included in an oligo pool assay (OPA) investigating genetic variants from different candidate pathways. Data for an additional non-synonymous TLR4 SNP (D299G; rs4986790) that had been previously genotyped by TaqMan assay in the NCI-SEER (14) and Connecticut studies were also included in the pooled analysis, increasing the total number of investigated SNPs to 37. Gene coverage (the number of genotyped SNPs divided by the number of SNPs from the design set of SNPs in Hamap Build 20 CEU samples) was calculated to be 61, 69 and 94% for TLR10–TLR5, TLR6, TLR2 and TLR4, respectively. All investigated SNPs in TLR1, TLR2, TLR4, TLR6 and TLR10 had genotype concordance rates and assay completion rates >99%. For each SNP, an exact test of Hardy–Weinberg proportions was performed among Caucasian controls. One TLR2 SNP deviated significantly from Hardy–Weinberg proportions ($t^{2} = 4666187, P = 0.02$).

We excluded samples with OPA-wide sample completion rates <90% (NCI-SEER: 11 cases, 6 controls; Connecticut: 2 controls; NSW: 4 cases, 9 controls). The final analytic population from pooling the three studies consisted of 1946 cases and 1808 controls (NCI-SEER, 990/828; Connecticut, 436/515; NSW, 520/465).

**Statistical methods**
Statistical analyses were performed using SAS version 9.1 (SAS Institute, Cary, NC) unless otherwise specified. Odds ratios (ORs) and 95% confidence intervals (CIs) estimating the relative risk of NHL in relation to SNP genotype were calculated using unconditional logistic regression adjusting for age, race/ethnicity, sex and study center. Analyses restricted to non-Hispanic Caucasians were also performed. Tests for trend (i.e. assuming a codominant genetic model) were performed by modeling the number of rare alleles (0, 1 or 2) as
a continuous variable. In general, \( P \)-values for SNPs reported in the Results are from the trend test. For rs4833103, which showed consistent evidence of association under a dominant genetic model throughout our analyses, both trend test \( P \)-values and \( P \)-values assuming a dominant model are reported.

Gene-level statistical tests of association of \( TLR10-TLR1-TLR6 \), \( TLR2 \) and \( TLR4 \) were conducted using two approaches. First, we performed the MinP test, which assesses the true statistical significance of the smallest \( P \)-trend within each gene region by permutation-based resampling methods (10 000 permutations) that automatically adjust for the number of tag SNPs tested within that gene and the underlying linkage disequilibrium (LD) pattern (25,26). These analyses were performed using the MATLAB Statistics Toolbox (http://www.broad.mit.edu/mpg/haploview/). For both approaches, codominant haplotypes were modeled with adjustment for age, sex and study center.

Two types of haplotype analyses were conducted among non-Hispanic Caucaisan subjects only. We evaluated risk associated with haplotypes defined by SNPs within a sliding window of three loci across a gene region (Haplo.Stats, haplo.glm, minimum haplotype frequency 1%). These region-wide haplotype analyses were performed after visualizing the haplotype structures across \( TLR10-TLR1-TLR6, TLR2 \) and \( TLR4 \) and noting the absence of separate haplotype blocks within each region (Haplovie version 4.0; http://www.broad.mit.edu/mpg/haplovie/). For both approaches, codominant haplotype effects were modeled with adjustment for age, sex and study center.

In addition to analyses of all NHL, we investigated associations with specific histologic subtypes (FL, DLBCL, MZL and CLL/SLL). Polytomous regression modeling was used to simultaneously calculate subtype-specific SNP effects and to construct tests of heterogeneity in codominant SNP effects across subtypes. Subtype-specific gene-level tests and haplotype analyses were performed using unconditional logistic regression modeling. The association between the \( TLR2 \) variant rs3804100 and MALT lymphoma, the most common type of MZL, was also investigated using unconditional logistic regression.

Results

The distributions of cases and controls were similar with respect to sex, age, race or ethnicity and study site, both overall and within each study (Table II). Approximately 90% of the overall study population were non-Hispanic Caucasian (87% of controls and 90% of cases), 5% were Black and 6% were members of other racial/ethnic groups. The NCI-SEER and Yale studies had similar distributions of NHL subtypes, whereas the NSW study had a higher frequency of FL (37 versus 24%) and a lower frequency of CLL/SLL (3 versus 10%) and not otherwise specified (NOS) tumors (7 versus 15%) compared with the other studies.

Figure 1 summarizes findings from tests of association involving the 11 investigated SNPs within the \( TLR10-TLR1-TLR6 \) gene cluster. Two SNPs were associated with NHL risk: rs10008492 (OR\(_{CT}\) 1.12, 95% CI 0.97–1.30; OR\(_{TT}\) 1.40, 95% CI 1.15–1.71; \( P \)-trend = 0.001) and rs4833103 (OR\(_{AC}\) 0.75, 95% CI 0.64–0.88; OR\(_{AA}\) 0.74, 95% CI 0.62–0.90; \( P \)-trend = 0.002; \( P \)\(_{dominant}\) = 0.0002). When analyses were restricted to non-Hispanic Caucasians, both SNPs remained associated with NHL (rs10008492, \( P \)-trend = 0.008; rs4833103, \( P \)-trend = 0.01, \( P \)\(_{dominant}\) = 0.003). Associations with these SNPs were

<table>
<thead>
<tr>
<th>Table II. Study-specific and pooled demographic and pathology characteristics of study participants in the NCI-SEER, Yale and NSW NHL case–control studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NCI</strong></td>
</tr>
<tr>
<td>Control, ( N = 828 )</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
</tr>
<tr>
<td>&lt;50</td>
</tr>
<tr>
<td>50–59</td>
</tr>
<tr>
<td>60–69</td>
</tr>
<tr>
<td>70+</td>
</tr>
<tr>
<td><strong>Race/ethnicity</strong></td>
</tr>
<tr>
<td>White, non-Hispanic</td>
</tr>
<tr>
<td>Black</td>
</tr>
<tr>
<td>Other/unknown</td>
</tr>
<tr>
<td><strong>Study site</strong></td>
</tr>
<tr>
<td>Detroit</td>
</tr>
<tr>
<td>Iowa</td>
</tr>
<tr>
<td>Los Angeles</td>
</tr>
<tr>
<td>Seattle</td>
</tr>
<tr>
<td>Connecticut</td>
</tr>
<tr>
<td>NSW</td>
</tr>
<tr>
<td><strong>ACT</strong></td>
</tr>
<tr>
<td><strong>NHL subtype</strong></td>
</tr>
<tr>
<td>DLBCL</td>
</tr>
<tr>
<td>Follicular</td>
</tr>
<tr>
<td>CLL/SLL</td>
</tr>
<tr>
<td>Mantle cell</td>
</tr>
<tr>
<td>Marginal zone</td>
</tr>
<tr>
<td>LPL</td>
</tr>
<tr>
<td>MF/SS</td>
</tr>
<tr>
<td>Burkitt</td>
</tr>
<tr>
<td>Peripheral T</td>
</tr>
<tr>
<td>NOS</td>
</tr>
<tr>
<td><strong>DNA source</strong></td>
</tr>
<tr>
<td>Blood</td>
</tr>
<tr>
<td>Buccal</td>
</tr>
</tbody>
</table>

ACT, Australian Capital Territory; LPL, lymphoplasmacytic lymphoma; MF/SS, mycosis fungoides/sézary syndrome.
SNPs are in supplementary Tables II–VI, available at Carcinogenesis Online). Only some subtype-specific associations were statistically significant (FL for rs10008492; DLBCL and MZL for rs4833103). However, for each SNP, the associations were in the same direction for all subtypes. Subtype-specific sliding-window haplotype analyses (supplementary Table VII is available at Carcinogenesis Online) and global tests of association (supplementary Table VIII is available at Carcinogenesis Online) of the TLR10–TLR1–TLR6 region supported associations with DLBCL and, less consistently, FL and CLL/SLL.

Investigations of the 11 tag SNPs summarizing the TLR2 region did not provide strong evidence of association with NHL risk (supplementary Table I is available at Carcinogenesis Online), although two tag SNPs, located within the neighboring gene RNF175, were associated with risk at a moderate level of statistical significance (rs11935252, \( P_{\text{hap}} = 0.01 \); rs7695605, \( P_{\text{hap}} = 0.03 \)). Of the global tests of TLR2 variation performed, the likelihood ratio test was statistically significant \( (P = 0.005) \), but the MinP test and region-wide haplotype global test were not \( (P = 0.12 \) and 0.84, respectively).

While subtype-specific analyses of TLR2 SNPs were generally null, one notable exception was observed for rs3804100 (Table III), which was strongly associated with MZL risk (CT/CC versus TT: \( OR = 1.89, 95\% \ CI 1.27–2.81; P = 0.002 \)). Statistically significant associations with this SNP were observed in two of the three studies (Figure 2). Analyses of rs3804100 among non-Hispanic Caucasians only yielded identical findings \( (P = 0.001) \), and sliding-window haplotype analysis also supported an association with this SNP (supplementary Table VII is available at Carcinogenesis Online). Additional analyses restricted to MALT lymphoma \( (n = 118) \), the most common type of MZL, yielded an even stronger association \( (OR_{\text{CT/CC}} = 2.21, 95\% \ CI 1.42–3.43; P = 0.0005) \). This variant was not associated with other subtypes or with NHL risk overall; a test of heterogeneity in SNP effects across subtypes was statistically significant \( (P = 0.007) \). The MinP test of TLR2 variants approached statistical significance for MZL \( (P = 0.06) \), but the likelihood ratio test and region-wide haplotype global test did not \( (P = 0.15 \) and 0.14, respectively).

The 15 investigated SNPs in TLR4 were not associated with NHL (supplementary Table I is available at Carcinogenesis Online). Gene region tests, haplotype-based analyses and investigations by subtype were similarly null.

Discussion

In this pooled investigation of TLR gene polymorphisms in three case–control studies of NHL, we found consistent evidence that variation in the TLR10–TLR1–TLR6 region is associated with disease risk. Global tests of association across all genotyped variants in this region were statistically significant. In particular, two SNPs within the region, rs10008492 and rs4833103, were significantly associated with NHL; these SNP associations were observed in all three studies and did not appreciably differ by histologic subtype. We did not observe clear evidence of association with variation in TLR2, although the variant rs3804100 was significantly associated with MZL. No associations with TLR4 variants were observed.

TLR10, TLR1 and TLR6 are located in a gene cluster spanning 57 kb on chromosome 4p14. TLR-1 and -6 are important elements of TLR-2 signaling; these polypeptides, which bind tri-acyl and di-acyl lipoproteins, respectively, interact with TLR-2 to create heterodimeric receptors capable of recognizing a broad spectrum of pathogen ligands (2). The ligand and function of TLR-10 are not known, although this receptor has been shown to be expressed in normal and malignant B lymphocytes (8). There is evidence that genetic variation in the TLR10–TLR1–TLR6 region may influence risk of inflammatory diseases; associations with asthma (10,11) and aspergillosis following allogeneic stem cell transplantation (28) have been reported. Associations with TLR10–TLR1–TLR6 variants were also observed in a large case–control study of prostate cancer (12), although a subsequent study of this region reported null findings (29).
The functional relevance of the tag SNPs rs10008492, located 8.5 kb telomeric of TLR10, and rs4833103, located 9.1 kb centromeric of TLR1 and 12.9 kb telomeric of TLR6, has not been investigated. If our observed associations are real, it is probable that these two variants do not directly influence NHL susceptibility but rather are markers for one or more underlying causal variants. Given the strong LD throughout the TLR10–TLR1–TLR6 region (Figure 1), it is difficult to infer a specific effect of any one of these genes from these SNP findings. It is worth noting though that, within the HapMap CEU sample, both variants exhibit LD with the non-synonymous SNP I602S, which has been shown to exhibit reduced activity toward tri-acyl lipoproteins in two studies (30,31). It is possible that I602S and/or other functionally relevant TLR1 variants such as P315L (32) directly influence NHL risk. However, we cannot rule out TLR10 and TLR6 as susceptibility loci; indeed, a rare non-synonymous SNP was significantly associated with NHL risk in a recent case–control analysis (9). Additional investigations involving a fine-mapping approach to TLR10–TLR1–TLR6 will be important for identifying the underlying causal variants in this region.

We did not observe clear evidence of association between TLR2 variants and risk of all NHL; the likelihood ratio test was statistically significant, but findings from other analyses were null or equivocal. The TLR2 region tag SNPs rs11935252 and rs7695605 were associated with NHL at a moderate level of statistical significance; however, as both variants actually reside within the neighboring gene RNF175, involved in metal ion binding, it is unclear what gene effect would be responsible for those associations (if real). We did observe a strong association with MZL risk for the variant rs3804100, both overall and within two of the three studies. No TLR2 variants were associated with other NHL subtypes. Nieters et al. (6) reported a statistically significant association with FL for the TLR2 –16933T>A (rs4696480) variant. This SNP was not genotyped in our study, although we note that it resides within a LD block shared by three SNPs included in our study (rs6835636, rs4696187 and rs13150331) in the HapMap CEU sample.

TLR2 haplotypes including rs3804100 have been previously associated with risk of type 1 diabetes and severity of genital herpes simplex virus type 2 infection (33,34). If our observed association with rs3804100 is real, it is most probably that this SNP, which is synonymous (S450S), is a marker for another causal variant. Several non-synonymous SNPs within the TLR2 exon have been identified, including the putatively functional variant Arg753Gln (rs5743708) (35–37), although most are rare. An association between TLR2 and MZL is biologically plausible given the strong evidence linking specific infectious organisms to the pathogenesis of MZL, MALT lymphoma in particular (38), and the importance of TLR-2 in recognizing these organisms. There is convincing evidence that H.pylori infection is a causal factor for gastric MALT lymphoma, the most common type of extranodal MZL (39,40). Other infectious organisms have also been linked to MALT lymphomas at other anatomic sites, including Borrelia burgdorferi (41,42), Chlamydia psittaci (43), Campylobacter jejuni (44) and the hepatitis C virus (45,46). Interestingly, TLR-2 has been shown to play an important role in mediating immune responses to H.pylori infection (47,48), B.burgdorferi (49) and hepatitis C virus (50).

Genetic variation in TLR4 was not associated with NHL risk in our study. There is conflicting evidence from smaller studies regarding an association of the TLR4 polymorphism Asp299Gly (rs4986790) with NHL; this variant was associated with increased risk of MALT lymphoma in one study (6) and with decreased risk of MALT lymphoma (7) and DLBCL (5) in two others. Our findings for this SNP were consistently null, both overall and for specific subtypes.

Strengths of this pooled analysis include its large size and the population-based design of the three participating case–control studies. It is unlikely that our findings are the result of bias due to population stratification, as race was adjusted for in regression modeling, and analyses restricted to non-Hispanic Caucasians yielded virtually identical findings. Given that we investigated several variants across the three gene regions for both all NHL and different subtypes, we must consider the possibility that our observed associations are

---

**Fig. 2.** Study-specific associations with NHL for rs10008492 and rs4833103 (TLR10–TLR1–TLR6) and with MZL for rs3804100 (TLR2). Square symbols represent ORs; symbol size is proportional to number of cases. Horizontal lines represent 95% CIs.
false-positive findings. The consistency of our key findings across the three participating studies suggests that they are not due to chance; however, these results require replication in other studies before meaningful inferences regarding causation are drawn.

In conclusion, this pooled investigation of TLR gene variants across the three case–control studies of NHL provides strong evidence that variation in the TLR10–TLR1–TLR6 region is associated with risk and suggests that TLR2 variants may influence susceptibility to MZL. Additional studies are needed to replicate these findings and, more generally, to explore further the relevance of TLR pathways to the pathogenesis of NHL. Pooled investigations within the InterLymph Consortium (51) will be especially informative in this regard.

Supplementary material
Supplementary Tables I–VIII can be found at http://carcin.oxfordjournals.org/

Funding
Intramural Research Program of the National Institutes of Health (National Cancer Institute) to all genotyping and statistical analysis for this work and NCI-SEER study; Public Health Service (N01-PC-65064, N01-PC-67008, N01-PC-67009, N01-PC-67010, N02-PC-71105) to NCI-SEER study; National Cancer Institute National Institutes of Health grant (CA62006) to Connecticut study; National Cancer Institute National Institutes of Health grant (CA62006, N01-PC-67008, N01-PC-67009, N01-PC-67010, N02-PC-71105) to NSW Central Cancer Registry.

Acknowledgements
We thank Mary McAdams, Peter Hui, Michael Stagner and Zeynep Kalaylioglu of Information Management Services for their programming support. For the NCI-SEER study, we also gratefully acknowledge the contributions of the staff and scientists at the SEER centers of Iowa, Los Angeles, Detroit and Seattle for the conduct of the study’s field effort. The NSW study was made possible by access to new notifications to the NSW Central Cancer Registry. Ann-Maree Hughes oversaw conduct of the study and Melissa Litchfield, Maria Agaliotis, Chris Goumas, Jackie Turner and staff of the Hunter Valley Research Foundation contributed to the data collection. Jenny Turner, study pathologist, reviewed all pathology reports and original slides as necessary.

Conflict of Interest Statement: None declared.

References


Received August 13, 2008; revised October 16, 2008; accepted November 18, 2008