Genetic variation in caspase genes and risk of non-Hodgkin lymphoma: A pooled analysis of three population-based case-control studies

Qing Lan¹, Lindsay M. Morton¹, Bruce Armstrong², Patricia Hartge¹, Idan Menashe¹, Tongzhang Zheng³, Mark P. Purdue¹, James R. Cerhan⁴, Yawei Zhang³, Andrew Grulich⁵, Wendy Cozen⁶, Meredith Yeager⁷, Theodore R. Holford³, Claire M. Vajdic⁸, Scott Davis⁹, Brian Leaderer³, Anne Kricker², Maryjean Schenk¹⁰, Shelia H Zahm¹, Nilanjan Chatterjee¹, Stephen J. Chanock¹,⁷, Nathaniel Rothman¹, Sophia S. Wang¹.

¹ Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS, Rockville, MD
² School of Public Health, The University of Sydney, Sydney, New South Wales, Australia
³ Department of Epidemiology and Public Health, Yale School of Medicine, New Haven, CT
⁴ Mayo Clinic, College of Medicine, Rochester, MN
⁵ National Centre in HIV Epidemiology and Clinical Research, University of New South Wales, Sydney, New South Wales, Australia
⁶ Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA
⁷ Core Genotyping Facility, Advanced Technology Center, National Cancer Institute, NIH, DHHS, Gaithersburg, MD
⁸ UNSW Cancer Research Centre, Prince of Wales Clinical School, University of New South Wales, Sydney, Australia
⁹ Fred Hutchinson Cancer Research Center and University of Washington, Seattle, WA
¹⁰ Department of Family Medicine and Karmanos Cancer Institute, Wayne State University, Detroit, MI

Running Title: Caspase gene polymorphisms and non-Hodgkin lymphoma

Key Words: non-Hodgkin lymphoma, Single nucleotide polymorphisms, Caspase, Apoptosis

Author for correspondence: Qing Lan, M.D., Ph.D., Occupational & Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS, MSC 7240, 6120 Executive Blvd., EPS 8109, Bethesda, Maryland 20892-7240

Telephone: 301-435-4706
Fax: 301-402-1819
E-mail: qingl@mail.nih.gov
Abstract

Caspases play a critical role in regulation of apoptosis, cell differentiation, inflammation, and innate immunity, and several are mutated or have altered expression in non-Hodgkin lymphoma (NHL). To study the impact of genetic variation in caspases on NHL risk, we analyzed tag single nucleotide polymorphisms (SNPs) in 12 caspase and related genes in three population-based case-control studies (1,946 cases and 1,808 controls). Gene-based analysis, adjusting for the number of tagSNPs genotyped in each gene, showed significant associations for CASP8, CASP9, and CASP1. SNP-based analysis showed that CASP8 rs6736233 (odds ratio (OR) CG=1.21; ORCC=2.13; p-trend=0.011); CASP9 rs4661636 (ORCT=0.89; ORTT=0.77; p-trend=0.014); and CASP1 rs1785882 (ORAT=1.12; ORAA=1.30; p-trend=0.0054) were significantly associated with NHL risk and consistent across studies. It is noteworthy that genetic variants in CASP8 were associated with risk of all major NHL subtypes. Our findings suggest that genetic variation in caspases may play an important role in lymphomagenesis.
Introduction

Caspases are highly conserved intracellular cysteine proteases that mediate apoptosis, and are categorized as initiator caspases (1, 2, 4, 5, 8, 9, 10, 11, 12) or effector caspases (3, 6, 7, 14).\(^1\) Initiator caspases are the first to be activated in apoptosis and in turn activate effector caspases, orchestrating programmed cell death.\(^1\) Caspases contribute to biologic processes important in lymphomagenesis, including cytokine maturation (e.g., caspase 1, 5, 11), NF-κB activation (e.g., caspase 1, 2, 8), and B cell maturation/proliferation (e.g., caspase 3, 8).\(^1\)

There is growing evidence that caspase genes are altered in non-Hodgkin lymphoma (NHL).\(^2,3\) Somatic mutations in caspase 3 and 10 have been reported\(^3,4\) and NHL tumor expression array analyses have shown that caspase 1, 2, 9, and 10 were differentially expressed by NHL subtypes.\(^2,5,6\) There is also preliminary evidence from association studies that single nucleotide polymorphisms (SNPs) in a number of caspase and caspase-related genes may be associated with risk of NHL or its subtypes.\(^7,8\) To investigate whether genetic variation in caspase genes plays in lymphomagenesis, we genotyped tagSNPs in 12 caspase or caspase-related genes in 1,946 NHL cases and 1,808 controls pooled from three independent population-based case-control studies conducted in the US and Australia.

Materials and methods

Three population-based case-control studies of NHL participated in this pooled analysis: the NCI-SEER NHL study, conducted within the SEER Iowa, Detroit, Los Angeles and Seattle
registries; the Connecticut NHL study, conducted among female residents of Connecticut; and the NSW study, conducted among residents of New South Wales and the Australian Capital Territory, Australia (Supplementary Table 1, Supplementary materials). The protocols for each study were approved by the Institutional Review Boards of the NCI, each SEER center for the NCI-SEER study (Yale University, the Connecticut Department of Public Health, and the NCI for the Connecticut study), and all participating institutions for the NSW study. All study participants provided informed consent, in accordance with the Declaration of Helsinki. NHL subtypes were grouped according to the World Health Organization classification using the International Lymphoma Epidemiology Consortium (InterLymph) guidelines. DNA was extracted from blood or buccal cells, and tagSNPs were genotyped at the NCI Core Genotyping Facility (Supplementary materials). In total, 79 tagSNPs in 12 caspase or caspase-related genes and 23 additional SNPs within 9 regions adjacent to these genes used to expand genomic coverage were selected (Supplementary Table 2); the latter 23 SNPs did not show any noteworthy evidence of association (Supplementary Table 3). SNPs in caspase genes showing association in a previous report from the Connecticut study were completely tagged (Supplementary Table 2) and results for those SNPs in the pooled analyses can found in Supplementary Tables 3-8. However, stronger effects were found for novel SNPs genotyped in the same caspase genes for the pooled analysis reported here and are featured in the Results and Discussion section.

Genotype-specific risks of NHL for each SNP were estimated as odds ratios (ORs) and 95% confidence intervals (CI) for the heterozygote and less common homozygote genotypes, with the more common homozygote as the reference, using unconditional logistic regression. Our
approach for defining the referent group follows genetic convention for tagSNPs, but we note that the directionality of effects is to be determined by follow-up studies to establish the causal variants. Models were adjusted for age, race, sex and study center. Polytomous multivariate unconditional logistic regression models were used to evaluate the effect among different NHL subtypes.

To obtain a gene-level summary of association and adjust for the number of tag SNPs in each gene, taking into account the underlying linkage disequilibrium pattern, we computed the minimum p-value (“minP test”), which assesses the statistical significance of the smallest p-trend within each gene by permutation-based re-sampling methods (10,000 permutations). A significance level of $p<0.05$ was interpreted as evidence of association. To account for multiple comparisons across the 12 caspase genes tested, we applied the false discovery rate (FDR) to the minP test for all NHL. A FDR value $<0.2$ was considered evidence that an association had a relatively low probability of being a false discovery. We emphasize SNPs with associations for NHL risk overall, as well as with histologic subtypes, in contrast to SNPs that show association with only specific histologies, because the latter analyses have reduced power and are more likely to be false positive associations.

Haplotype analysis was carried out using an expectation-maximization (EM) algorithm and HaploStats (R version 1.2.0) among non-Hispanic Caucasians, but did not reveal additional insights beyond those obtained from the gene and region-based analyses (data not shown).
Results and Discussion

Cases and controls showed comparable distributions by age and race within each study (Supplementary Table 1). We found significant evidence of association at the gene-level for CASP8, CASP9, and CASP1 with NHL and one or more subtypes (Table 1, Figure 1). FDR values for the associations with NHL were 0.15. There was also evidence for association at the gene-level for CASP8AP2 and CASP14 with diffuse large B-cell lymphoma (DLBCL), CASP4 with chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL), and CASP2 with marginal zone lymphoma (Table 1). Within genes showing an association with NHL, CASP8 rs6736233 and rs3769821 ($r^2=0.13$, $D’=0.98$) and CASP1 rs1785882 were associated with increased risk and CASP9 rs4661636 was associated with decreased risk of NHL overall and with one or more of its subtypes (Figure 1, Supplementary Tables 3,4). These associations were consistent across study, age, sex, and among non-Hispanic Caucasians (Figure 1, Supplementary Tables 3, 5-9). It is also noteworthy that in the aggregate genetic variants in CASP8 were significantly associated with risk of all four major B cell subtypes (Figure 1, Supplementary Table 4). In addition, CASP1 rs1785882 was associated with risk of NHL overall and two of its subtypes (DLBCL and CLL/SLL) (Figure 1, Supplementary Table 4).

This is the first comprehensive evaluation of genetic variation in caspase genes and risk of NHL. Our results suggest that SNPs in initiator caspases (i.e., CASP8, CASP9, and CASP1) affect lymphomagenesis.
The main initiator caspases in mammals are caspase 8 (located in the death receptor-mediated apoptosis pathway) and caspase 9 (located in the intrinsic mitochondrial apoptosis pathway).\textsuperscript{1} Lan et al.\textsuperscript{19} recently reported that higher mitochondrial DNA copy number was associated with risk of NHL, which is consistent with impaired mitochondrial apoptosis. We acknowledge, however, that the observed associations could reflect other biologic functions mediated by these genes in addition to apoptosis. In particular, caspase 8 plays a broad role in regulating lymphocyte homeostasis, NF-κB activation, and differentiation of monocytes into macrophages,\textsuperscript{1} all of potential relevance for NHL etiology.\textsuperscript{20} One report found that individuals with \textit{CASP8} mutations had decreased T, B, and NK cell activation and decreased lymphocyte apoptosis.\textsuperscript{21} Although \textit{CASP8} rs1045485 has been associated with risk of breast cancer, melanoma, and glioma,\textsuperscript{22-24} we did not detect an association with NHL.

Caspase 1, 4, and 5 play a key role in maturation of proinflammatory cytokines in cells infected by certain pathogens, and caspase 1 is the most efficient caspase in the process.\textsuperscript{25} For example, caspase 1, initially known as interleukin-1B-converting enzyme, is critical for maturation of IL1B,\textsuperscript{1} and also regulates IFN-gamma production.\textsuperscript{25} Further, caspase 1 plays a role in NF-κB activation,\textsuperscript{1} so the underlying biologic basis of the \textit{CASPI} association we report here could be due to one or more of these functions.

In summary, our study provides evidence that common genetic variants in \textit{CASP8}, \textit{CASPI}, and \textit{CASPI} are associated with risk of NHL and one or more subtypes. If replicated in larger studies, a comprehensive strategy of fine mapping followed by functional analyses should be carried out and gene-environment interactions should be explored.
Acknowledgements

DNA extraction, genotyping and statistical analysis for this project were supported by the Intramural Research Program of the National Institutes of Health (NIH) (National Cancer Institute (NCI)). The NCI-SEER study was also supported by the Intramural Research Program of the NIH (NCI), and by Public Health Service (PHS) contracts N01-PC-65064, N01-PC-67008, N01-PC-67009, N01-PC-67010, and N02-PC-71105. The Connecticut study was supported by NIH grant CA62006 (TZ) from the NCI. The NSW study was supported by the National Health and Medical Research Council of Australia Project Grant number 990920 (BA), The Cancer Council NSW, and The University of Sydney Medical Foundation. We gratefully acknowledge the assistance of Peter Hui (Information Management Services, Inc., Silver Spring, MD) for programming support.

Authorship

CM, Davis S, Leaderer B, Kricker A, Schenk M, Zahm SH, Chatterjee N, and Chanock SJ. The authors have no conflicts of interest to declare.

Reference List


<table>
<thead>
<tr>
<th>Genes</th>
<th>SNPs /gene #</th>
<th>NHL P value</th>
<th>DLBCL P value</th>
<th>Follicular P value</th>
<th>MZL P value</th>
<th>CLL/SLL P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASP1</td>
<td>4</td>
<td>0.019</td>
<td>0.032</td>
<td>0.48</td>
<td>0.74</td>
<td>0.0014</td>
</tr>
<tr>
<td>CASP2</td>
<td>4</td>
<td>0.18</td>
<td>0.34</td>
<td>0.34</td>
<td>0.023</td>
<td>0.76</td>
</tr>
<tr>
<td>CASP3</td>
<td>6</td>
<td>0.16</td>
<td>0.46</td>
<td>0.62</td>
<td>0.74</td>
<td>0.43</td>
</tr>
<tr>
<td>CASP4</td>
<td>6</td>
<td>0.49</td>
<td>0.94</td>
<td>0.37</td>
<td>0.42</td>
<td>0.01</td>
</tr>
<tr>
<td>CASP5</td>
<td>10</td>
<td>0.50</td>
<td>0.18</td>
<td>0.61</td>
<td>0.86</td>
<td>0.08</td>
</tr>
<tr>
<td>CASP6</td>
<td>5</td>
<td>0.37</td>
<td>0.25</td>
<td>0.71</td>
<td>0.37</td>
<td>0.23</td>
</tr>
<tr>
<td>CASP7</td>
<td>13</td>
<td>0.73</td>
<td>0.82</td>
<td>0.66</td>
<td>0.37</td>
<td>0.87</td>
</tr>
<tr>
<td>CASP8</td>
<td>9</td>
<td>0.027</td>
<td>0.16</td>
<td>0.050</td>
<td>0.033</td>
<td>0.042</td>
</tr>
<tr>
<td>CASP9</td>
<td>4</td>
<td>0.036</td>
<td>0.22</td>
<td>0.043</td>
<td>0.23</td>
<td>0.45</td>
</tr>
<tr>
<td>CASP10</td>
<td>3</td>
<td>0.05</td>
<td>0.35</td>
<td>0.21</td>
<td>0.13</td>
<td>0.06</td>
</tr>
<tr>
<td>CASP14</td>
<td>9</td>
<td>0.60</td>
<td>0.040</td>
<td>0.83</td>
<td>0.21</td>
<td>0.27</td>
</tr>
<tr>
<td>CASP14</td>
<td>6</td>
<td>0.31</td>
<td>0.0073</td>
<td>0.57</td>
<td>0.18</td>
<td>0.16</td>
</tr>
</tbody>
</table>

* Abbreviations: NHL, non-Hodgkin lymphoma; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; CLL/SLL, chronic lymphocytic leukemia/small lymphocytic lymphoma; MZL, marginal zone lymphoma. Bold type indicates p<0.05. The minP test assesses the true statistical significance of the smallest p-trend within each gene (determined by dichotomous logistic regression, comparing NHL or NHL subtypes to controls; SNPs listed in Supplementary Table 2) 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 pattern.13

1 CASP1: significant SNPs for NHL, rs1785882; DLBCL, rs1785882; CLL/SLL, rs1785882, rs501626.
2 CASP2: significant SNPs for MZL, rs7810486.
3 CASP3: significant SNPs for FL, rs78791740, rs11226565, rs7123277.
4 CASP4: significant SNPs for NHL, rs6736233, rs3769821; Follicular, rs6736233, rs3769821, rs2293554; MZL, rs3769825, rs3769821, rs700636; CLL/SLL, rs3769825, rs6736233, rs3769821.
5 CASP9: significant SNPs for NHL, rs4661636, rs4646047.
6 CASP8AP2: significant SNPs for DLBCL, rs12661230.
7 CASP14: significant SNPs for DLBCL, rs714920.
Figure 1. Association between the most noteworthy SNPs in CASP8, CASP9, and CASP1 and risk of NHL by study, NHL subtype, sex, age, and ethnicity, based on the additive model. Square symbols represent odds ratios (OR); symbol size is proportional to number of cases. Horizontal lines represent 95% confidence intervals. Number of cases and controls by study - NCI-SEER (990 cases, 828 controls), Connecticut (436 cases, 515 controls), New South Wales (520 cases, 465 controls); by NHL subtype - DLBCL (Diffuse large B-cell lymphoma) n = 600, Follicular, n = 540, Marginal Zone n = 160, and CLL/SLL (chronic lymphocytic leukemia/small lymphocytic lymphoma) n = 161; by sex – males (840 cases, 711 controls), females (1106 cases, 1097 controls); by age - < 50 (484 cases, 408 controls), 50+ (1462 cases, 1400 controls); by ethnicity - Non-Hispanic Caucasians (1751 cases, 1578 controls), all ethnicities combined (1946 cases, 1808 controls). p-values are from additive (i.e., trend) model. MAF represents minor allele frequency.
Genetic variation in caspase genes and risk of non-Hodgkin lymphoma: a pooled analysis of three population-based case-control studies

Qing Lan, Lindsay M. Morton, Bruce Armstrong, Patricia Hartge, Idan Menashe, Tongzhang Zheng, Mark P. Purdue, James R. Cerhan, Yawei Zhang, Andrew Grulich, Wendy Cozen, Meredith Yeager, Theodore R. Holford, Claire M. Vajdic, Scott Davis, Brian Leaderer, Anne Kricker, Maryjean Schenk, Shelia H. Zahm, Nilanjan Chatterjee, Stephen J. Chanock, Nathaniel Rothman and Sophia S. Wang