Single nucleotide polymorphisms in the mitochondrial displacement loop and age at onset of non-Hodgkin lymphoma

Haiyan Fan¹
Cuiju Wang²
Zhanjun Guo¹

¹Department of Gastroenterology and Hepatology, ²Department of Gynaecology Ultrasound, The Fourth Hospital of Hebei Medical University, Shijiazhuang, People’s Republic of China

Objective: Single nucleotide polymorphisms (SNPs) accumulated frequently in the mitochondrial displacement loop (D-loop) in many cancers. We had identified cancer risk-associated SNPs in the D-loop of non-Hodgkin lymphoma (NHL) patients previously, in this study, we investigated the association of age at onset and D-loop SNPs in NHL patients.

Materials and methods: The D-loop region of mtDNA was sequenced for 133 NHL patients recorded at the Fourth Hospital of Hebei Medical University. The Kaplan–Meier method was used to identify age at onset-associated SNPs in the D-loop of NHL patients. The Cox proportional hazards model was used to identify independent risk factors for age at onset.

Results: The SNP sites of nucleotides 146C/T, 151T/C, 194T/C, 315C/C insert, 523Del/A, and 525Del/C were identified for their association with age at onset, by the logrank test. In an overall multivariate analysis, allele 146 (relative risk, 0.403; 95% confidence interval [CI]: 0.182–0.895) (P = 0.026), allele 151 (relative risk, 0.378; 95% CI: 0.165–0.868) (P = 0.022), and allele 315 (relative risk, 3.554; 95% CI: 1.344–9.400) (P = 0.011) were identified as independent predictors for age at onset in NHL patients.

Conclusion: SNPs in the D-loop can predict age at onset in NHL patients. Analysis of the D-loop SNPs can help identify NHL patient subgroups at high risk of early onset.

Keywords: Kaplan–Meier; Cox proportional hazards model, D-loop, single nucleotide polymorphisms

Introduction
Non-Hodgkin lymphoma (NHL) is the fifth most frequent form of cancer among men and women, accounting for 2.15% of all cancer-related deaths in the People’s Republic of China over the period 2003–2007. Although the etiology of NHL is poorly understood, the immune-related conditions, family histology of hematopoietic malignancy, and body mass index have been shown to have an association with the cancer risk. Some studies have also shown that genetic factors are important to development of this cancer. The polymorphism of genes in the oxidative stress pathway also was identified for their association with a risk of NHL. However, the true mechanism of this cancer remains unknown. Although some studies have focused on genetics associated with the age at onset of NHL, they have yet to demonstrate a genetic basis of this disease.

The human mitochondrial genome is 16 kilobase (kb) in length and is formed as a closed-circular duplex molecule. Mitochondrial deoxyribonucleic acid (mtDNA) is believed to be more susceptible to DNA damage and so acquires mutations at a higher rate than nuclear DNA because of the high levels of reactive oxygen species (ROS)
generation, lack of protective histones, and limited DNA repair capacity in the mitochondria.\textsuperscript{13,14}

The mitochondrial displacement loop (D-loop) region is important for regulating both replication and expression of the mitochondrial genome because it contains the leading-strand origin of replication and the main promoter for transcription. In cancer patients, sequence changes accumulated extensively in this region.\textsuperscript{15–18} Previously, we sequenced the D-loop region in a case-control study of NHL patients and identified 140 single nucleotide polymorphisms (SNPs), including 26 SNPs with frequency distribution of minor allele greater than 5%, and we also identified NHL risk-associated SNPs.\textsuperscript{19}

In this study, we investigated the association of age at onset and SNPs of the D-loop, in NHL patients.

**Materials and methods**

**Tissue specimens and DNA extraction**

Diagnosis of lymphoma was made according to the WHO Classification of Tumors.\textsuperscript{20} Blood samples were collected from 133 NHL patients who received treatment in the Respiratory Department at the Fourth Hospital of Hebei University, Shijiazhuang, People’s Republic of China, between 1990 and 2011. The genomic deoxyribonucleic acid (DNA) was immediately extracted using the Wizard\textsuperscript{®} Genomic DNA Purification Kit (Promega Corporation, Fitchburg, WI, USA) and stored at $-20^\circ\text{C}$. All procedures were supervised and approved by the Human Tissue Research Committee at the hospital. Informed consent was obtained from all participants before enrollment.

**Polymerase chain reaction (PCR) amplification and sequence analysis**

A 982-base pair (bp) product of D-Loop was amplified with the forward primer 5’-CCCCATGCTTACAAGCAAGT-3’ (nucleotide 16190–16209) and reverse primer 5’-GCTTT-GAGGAGGTAAGCTAC-3’ (nucleotide 602–583). PCR was run with a PCR Master Mix Kit (Promega Corporation) and purified prior to sequencing. Cycle sequencing was performed with the BigDye Terminator v3.1 Cycle Sequencing Kit (Life Technologies, Carlsbad, CA, USA), and the products were then read with the ABI PRISM\textsuperscript{®} 3100 Genetic Analyzer (Life Technologies). Polymorphisms were confirmed by repeated analyses from both strands.

**Statistical analysis**

The age at onset curve of NHL patients was calculated, using the Kaplan–Meier method at each of the SNP sites, and compared with the logrank test. Multivariate analysis was performed with a Cox proportional hazards model. All of the statistical analysis was done with the SPSS Statistics for Windows, Version 17.0 software package (SPSS Inc, Chicago, IL, USA). A $P$-value of $<0.05$ was considered statistically significant.

**Results**

The age at onset distribution of the NHL patients was listed in Table 1. Those analyzed included: 25 patients at age less than 30 years old; 16 patients at age 30–40; 28 patients at age 41–50; 30 patients at age 51–60; 20 patients at age 61–70, and 14 patients at age older than 70. Age at onset was analyzed together with clinical characteristics, including gender and TNM (tumor node metastasis) classification, using the Kaplan–Meier method, and these were compared by the logrank test. TNM stage was associated with the age at onset, at statistically significant levels (Table 2 and Figure 1).

Twenty-six SNPs with minor allele frequency higher than 5% were subsequently evaluated for their association with age at onset. The NHL patients were divided into two groups on the basis of their genotype at each SNP site, and the association of SNPs in the D-loop and age at onset was calculated with Kaplan–Meier method. The SNP sites of nucleotides 146C/T, 151T/C, 194T/C, 315C/C insert, 523Del/A, and 525Del/C were identified for their association with age at onset by the logrank test (Table 3 and Figure 2); the SNP site of nucleotide 16261 was also identified for its association with age at onset but with borderline statistical significance ($P = 0.083$). We performed multivariate analysis for these predictors, including these SNPs and the clinical characteristics, with the Cox proportional hazards model. Allele 146 (relative risk, 0.403; 95% confidence interval [CI]: 0.182–0.895) ($P = 0.026$), allele 151 (relative risk, 0.378; 95% CI: 0.165–0.868) ($P = 0.022$), and allele 315 (relative risk, 3.554; 95% CI: 1.344–9.400) ($P = 0.011$) were identified for their association with age at onset.

**Table 1 Onset age distribution in NHL patients**

<table>
<thead>
<tr>
<th>Age</th>
<th>Cases (N)</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\leq 30$</td>
<td>20</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>30–40</td>
<td>5</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>41–50</td>
<td>16</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>51–60</td>
<td>20</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>61–70</td>
<td>10</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>$&gt;70$</td>
<td>10</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: NHL, non-Hodgkin lymphoma.
Table 2 Clinical characteristics and their association with onset age in NHL patients

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Cases (N)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>81</td>
<td>0.443</td>
</tr>
<tr>
<td>Female</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>TNM classification</td>
<td></td>
<td>0.031</td>
</tr>
<tr>
<td>I–II</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>III–IV</td>
<td>90</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: NHL, non-Hodgkin lymphoma; TNM, tumor node metastasis.

Table 3 Polymorphic sites of the D-loop and their association with onset age in NHL patients

<table>
<thead>
<tr>
<th>Polymorphic sites</th>
<th>Cases (N)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>146</td>
<td>C</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>0.006</td>
</tr>
<tr>
<td>151</td>
<td>T</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.024</td>
</tr>
<tr>
<td>194</td>
<td>T</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Note: including C and CC insertion.

Abbreviations: D-loop, displacement loop; NHL, non-Hodgkin lymphoma.

Discussion

Many SNPs in the D-loop region have been identified for their association with cancer risk and disease outcome in cancers.21–27 The present study has extended those analyses to determine the relationship between age at onset and germline SNPs, in NHL patients. The alleles 146, 151, and 315 were identified for their association with age at onset, at statistically significant levels, by multivariate analysis.

We have previously identified age at onset-associated SNPs of the D-loop, in patients with hepatocellular carcinoma and esophageal squamous carcinoma.28,29 In this study, we also identified age at onset-associated SNPs in NHL patients.

All of the cancer risk-associated SNP sites are located in the hypervariable segment (HVII) region, mutational hotspots at which germline and tumor mtDNA mutations preferentially occur.30 We have identified outcome-associated SNPs of these regions in other cancers.21,27 The allele 315 was identified as a cancer risk-associated SNP for NHL, and allele 151 was identified as a cancer risk-associated SNP for NHL patients (Table 4).
non–small cell lung cancer in previous study. Our data imply the important role of HVII region in modifying the onset age of cancer.

Since the mitochondrial D-loop is important for replication and expression of the mitochondrial genome, SNPs in this region might alter the function of electron transport chains by modifying mtDNA replication and transcription, which is responsible for the generation of ROS, and could contribute to nuclear genome damage as well as cancer initiation and promotion. These ROS-mediated mechanisms may accelerate earlier onset of this disease.

In conclusion, SNPs in the D-loop can predict age at onset in NHL patients, and analysis of D-loop SNPs can help to identify NHL patient subgroups at high risk of early onset.

**Table 4** Multivariate analysis of predictors associated with age at onset of NHL patients

<table>
<thead>
<tr>
<th>Factors</th>
<th>Relative risk</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>16261</td>
<td>0.714</td>
<td>0.329–1.548</td>
<td>0.393</td>
</tr>
<tr>
<td>146</td>
<td>0.403</td>
<td>0.182–0.895</td>
<td>0.026</td>
</tr>
<tr>
<td>151</td>
<td>0.378</td>
<td>0.165–0.868</td>
<td>0.022</td>
</tr>
<tr>
<td>194</td>
<td>0.515</td>
<td>0.192–1.381</td>
<td>0.187</td>
</tr>
<tr>
<td>315–316</td>
<td>3.554</td>
<td>1.344–9.400</td>
<td>0.011</td>
</tr>
<tr>
<td>523</td>
<td>1.330</td>
<td>0.392–4.515</td>
<td>0.648</td>
</tr>
<tr>
<td>525</td>
<td>0.400</td>
<td>0.121–1.322</td>
<td>0.133</td>
</tr>
<tr>
<td>TNM</td>
<td>1.601</td>
<td>0.927–2.764</td>
<td>0.091</td>
</tr>
</tbody>
</table>

**Abbreviations:** CI, confidence interval; NHL, non-Hodgkin lymphoma; TNM, tumor node metastasis.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**