Abstract. The accumulation of single nucleotide polymorphisms (SNPs) in the displacement loop (D-loop) of mitochondrial DNA (mtDNA) has been described for various types of cancer, and their association with cancer risk and disease outcome has been extensively identified. In the present study, we investigated the association between age-at-onset and SNPs in the mitochondrial D-loop using a population-based series of esophageal squamous cell carcinoma (ESCC) patients. The SNP sites of nucleotides 16266 C/T were identified for their association with age-at-onset using the log-rank test. The age-at-onset of patients with the minor allele T genotype was significantly lower than that of patients with the C genotype at the 16266 site (p=0.036). Genetic polymorphisms in the D-loop are predictive markers for age-at-onset in ESCC patients. Accordingly, the analysis of genetic polymorphisms in the mitochondrial D-loop may help identify ESCC patient subgroups at high risk of early onset.

Introduction

Esophageal cancer is one of the most common types of cancer in the population of Northern Central China with an age-standardized annual incidence rate >125/100,000 (1). Cumulative mortality attributed to esophageal cancer is approximately 20% for females and 25% for males (2). Despite improved diagnosis and therapeutic strategies, the prognosis of esophageal cancer remains poor, mostly due to its aggressive nature. Certain epidemiological factors were identified as risk factors or outcome predictors for ESCC (3,4), but the true mechanism of action for this cancer remains unknown. A few studies have focused on genetic factors associated with the age-at-onset of cancer. However, these studies have only demonstrated the genetic prevalence of this disease (4).

The human mitochondrial genome is 16 kb in length and is a closed-circular duplex molecule containing 37 genes, including 2 ribosomal RNAs and a complete set of 22 tRNAs (5). mtDNA is believed to be more susceptible to DNA damage and acquires mutations at a higher rate than nuclear DNA, due to high levels of reactive oxygen species (ROS), a lack of protective histones and a limited capacity for DNA repair in the mitochondria (6,7).

In cancer patients, sequence changes accumulate extensively in the mitochondrial D-loop region, which is significant for regulating the replication and expression of the mitochondrial genome, as it contains the leading-strand origin of replication and the main promoter for transcription (8-11). Using blood samples from the ESCC patients, we sequenced the D-loop, which contains a length of 1122 bp (nucleotides 16024-16569 and 1-576), and referred to the mitochondria database (www.mitomap.org). A total of 88 single nucleotide polymorphisms (SNPs) were identified in the D-loop. We also identified cancer outcome-associated SNPs using the Kaplan-Meier method (12). The aim of this study was to assess the correlation between germline SNPs of the D-loop and age-at-onset in ESCC patients.

Materials and methods

Tissue specimens and DNA extraction. Blood samples were collected at the Fourth Hospital of Hebei University from 68 patients who underwent esophageal cancer resection in the Department of Thoracic Surgery between 2003 and 2005. The patients were selected when they received endoscopy examination and specimens were confirmed as ESCC by a pathologist. The patients were from the Hebei Province of China, which is a high-risk area for ESCC. From the blood
samples, genomic DNA was immediately extracted using a Wizard Genomic DNA extraction kit (Promega, Madison, WI, USA). The study was approved by the Human Tissue Research Committee of the Fourth Hospital of Hebei Medical University. All patients provided written informed consent for the collection of samples and subsequent analysis.

**PCR amplification and sequence analysis.** The forward 5'-'CCCCCATGCTTACAAGCAAGT-3' (nucleotide 16190-16209) and reverse primers 5'-'GCTTTGAGGAGGTAA GCTAC-3' (nucleotide 602-583) were used for the amplification of a 982 bp product from the mtDNA D-loop region, as previously described (13). PCR was performed according to the protocol of the PCR Master Mix kit (Promega) and purified prior to sequencing. Cycle sequencing was carried out using the Dye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA, USA) and the products were then separated using the ABI PRISM Genetic Analyzer 3100 (Applied Biosystems). Polymorphisms were confirmed by repeated analyses from the two strands. SNPs were identified directly from the blood mitochondria.

**Statistical analysis.** The age-at-onset curve of ESCC patients was calculated using the Kaplan-Meier method at each SNP site and compared with the log-rank test. All of the statistical analysis was conducted using the SPSS 11.5 software package (SPSS Company, Chicago, IL, USA). P<0.05 was considered to indicate a statistically significant difference.

**Results**

**Age-at-onset distribution in ESCC patients and clinical characteristics.** A total of 68 patients who shared the same performance status (ECOG score, 0) were enrolled in this study. The age-at-onset distribution of the ESCC patients is listed in Table I. The patients analyzed included: 1 patient aged <40, 7 patients aged 40-50, 34 patients aged 51-60 and 26 patients aged >60 years old. None of these patients had received any adjuvant chemotherapy or radiation therapy following ESCC resection. The age-at-onset was analyzed against clinical characteristics using the Kaplan-Meier method and these were compared using the log-rank test. Gender, TNM stage and local lymphatic metastasis were not associated with age-at-onset at statistically significant levels (Table II).

**Correlation between mtDNA genotype and survival.** The correlation between the mtDNA genotype and survival was compared. SNPs with a minor allele frequency of <5% were included; thus, a remaining 22 SNPs were used for further analysis. The ESCC patients were divided into two groups based on their genotype at each SNP site. For each site, the age-at-onset curve was plotted using the Kaplan-Meier method for all ESCC patients. A marked difference in the age-at-onset appeared at nucleotide 16266, as revealed by the log-rank test (Fig. 1). This SNP was previously identified in the mitochondrial database (www.mitomap.org). The age-at-onset for patients with the minor allele 16266T genotype was significantly lower than that of patients with 16266C at the 16266 site (p=0.036) (Fig. 1).

**Discussion**

Selected SNPs in the D-loop region have been examined for their ability to predict cancer risk and outcome in other types of tumor (14-19). The present study has extended previous
analyses to determine the correlation between age-at-onset and germline SNPs in a continuous sequence of mtDNA between nucleotides 16190 and 583 in ESCC patients. The SNP 16266C/T was identified for its association with age-at-onset at a statistically significant level by the log-rank test.

We have performed a number of studies on D-loop SNPs in cancers of the digestive tract (18,19). In this study, we suggest for the first time that, apart from being a predictor for cancer risk and outcome, SNPs in the D-loop are also a predictor for age-at-onset in ESCC patients. Nucleotide 16266 is located between MT-5 and MT-3L on the mitochondrial genome. The functional significance of this SNP is not known, although we have found it to be associated with the risk of hepatocellular cell carcinoma by conducting case-control studies. Minor alleles of 16266T are associated with the early onset of ESCC; the age-at-onset curve decreased rapidly in patients carrying these alleles (Fig. 1). We compared the distribution frequency of this SNP between ESCC patients and normal controls. Among the 68 age-matched controls, only one carried the 16266T allele, whereas five ESCC patients carried the T allele. A much larger sample size is therefore required to perform a comparison for cancer risk at this site.

The D-loop region of mtDNA is significant for the regulation of mitochondrial genome replication and expression. SNPs in this region may affect mtDNA replication and lead to alteration of the electron transport chain, which is responsible for the release of highly reactive oxygen species (ROS), and may contribute to nuclear genome damage as well as cancer initiation and promotion (20-22). This SNP may alter the transcription of the mitochondrial genome, causing the production of ROS to be enhanced (23). This ROS-mediated mechanism may accelerate tumor development.

In conclusion, SNPs in the D-loop were found to be biomarkers for ESCC age-at-onset. The analysis of genetic polymorphisms in the D-loop may help to identify patient subgroups at a high risk of an early onset in ESCC cancers, thereby helping to refine therapeutic decisions.

Acknowledgements

This work was supported by the National Natural Science Foundation of P.R. China No. 30801384. The research was supported in part by the Natural Science Foundation of Hebei Province No. C2008000958.

References