Infection, Genetics and Evolution 42 (2016) 9–13

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

# The polymorphisms of *MSH6* gene are associated with AIDS progression in a northern Chinese population

Chuntao Wang [a,b,1], Chunyan Zhao [c,1], Xuelong Zhang [a], Lidan Xu [a], Xueyuan Jia [a], Haiming Sun [a], Jingcui Yu [d], Guangfa Zhang [a], Ning He [e], Qiuyan Li [a], Yuandong Qiao [a,*], Songbin Fu [a,*]

[a] Laboratory of Medical Genetics, Harbin Medical University, Harbin 150081, China
[b] Department of Biology, Mudanjiang Medical University, Mudanjiang 157011, China
[c] Department of Medical Biology and Genetics, Harbin Medical University-Daqing, Daqing 163319, China
[d] The Second Affiliated Hospital, Harbin Medical University, Harbin 150081, China
[e] Department of Clinical Laboratory, Qiqihar Traditional Chinese Medicine Hospital, Qiqihar 161000, China

ARTICLE INFO

*Article history:*
Received 19 February 2016
Received in revised form 12 April 2016
Accepted 14 April 2016
Available online 16 April 2016

*Keywords:*
*MSH6*
*MSH2*
SNPs
HIV-1
AIDS
Progression

ABSTRACT

It has been reported that DNA repair genes play an important role in HIV-1 infection and AIDS progression. One DNA repair pathway, the mismatch repair (MMR) is associated with a wide variety of tumors. However, the role of single nucleotide polymorphisms (SNPs) in the MMR genes and their importance in HIV-1 infection and AIDS progression remain unclear. In the present study, 479 HIV-1-infected and 487 healthy individuals from northern China were genotyped for nine SNPs in the *MSH2* gene (rs13019654, rs4608577, rs4952887, rs6726691, rs10191478, rs12999145, rs1981929, rs2042649, rs2303428) and five SNPs in the *MSH6* gene (rs2348244, rs3136245, rs3136329, rs2072447, rs7562048). Our results showed that the rs7562048 G allele frequency was significantly higher in the cases with the $CD4^+$ T-lymphocyte count <200 cells/μl than those with >200 cells/μl ($P = 0.001$, OR = 1.811, 95% CI 1.255–2.614), which is in agreement with the result of the Bonferroni correction. The frequencies of the rs2348244 C allele and rs3136245 T allele were higher in the cases at clinical phase IV than those at clinical phase I + II + III ($P = 0.026$, OR = 1.591, 95% CI 1.056–2.398 and $P = 0.019$, OR = 1.749, 95% CI 1.096–2.791, respectively); however, this difference is not supported by the Bonferroni correction. There were no significant differences in the frequency of allele, genotype and haplotype of the 14 SNPs between HIV-1-infected individuals and healthy controls ($P > 0.05$). These results suggest that the rs7562048 is associated with the clinical features and that the *MSH6* gene polymorphisms likely play an important role in the progression of AIDS in the northern Chinese population.



## 1. Introduction

Acquired immune deficiency syndrome (AIDS) is a significant and challenging chronic infectious disease resulting from human immunodeficiency virus (HIV) infection (Bao et al., 2012; Weiss, 1993). HIV infection leads to a progressive decline in the functionality and the number of $CD4^+$ T-lymphocyte, causing AIDS development (McCune, 2001). Studies have shown that there are a large quantity of apoptotic cells in the patients with HIV-1 infection, indicating that cell apoptosis is another mechanism responsible for the T-lymphocyte depletion in AIDS patients (Ameisen and Capron, 1991; Hel et al., 2006; Laurent-Crawford et al., 1991; Terai et al., 1991). One route of cell apoptosis is induced by DNA damage (De Flora et al., 1996), and thus the DNA repair system may play a role in HIV-1 infection and AIDS progression. It has been reported that DNA repair genes are associated with HIV-1 infection and AIDS progression. Specifically, *ERCC2* and *ERCC3* genes can promote viral cDNA replication and reduce the CD4 cell count by preventing degradation of retroviral cDNA (Yoder et al., 2006). The DNA-PK protein encoded by the *XRCC7* gene interacts with HIV-1 Tat to increase HIV-1 replication and transcription (Tyagi et al., 2011; Zhang et al., 2014). In addition, the polymorphisms of *XRCC1*, *XPG* and *ERCC2* genes may influence individual variations in the DNA repair capacity and play an important role in the progression of HIV diseases (Sobti et al., 2011; Sobti et al., 2010; Sobti et al., 2009).

The mismatch repair (MMR) is one of the DNA repair pathways, and its main function is to rectify base-pair mismatches to prevent gene mutation and maintain genomic stability and integrity (Burdett et al., 2001; Jiricny, 2006). The MMR pathway includes nine genes, of which, *MSH2* and *MSH6* are the two key components. The *MSH2* gene is located on chromosome 2p21–22, and required for recognizing nucleotide mismatches that occur during DNA replication (Lo et al., 2011). The *MSH6* gene is located on chromosome 2p16 and involved in the DNA base mismatch repair, reducing the spontaneous mutation rate (Suchy et al.,

* Corresponding authors at: 157 Baojian Road, Nangang District, Harbin 150081, China.
*E-mail addresses:* qiaoyuandong@hrbmu.edu.cn (Y. Qiao), fusb@ems.hrbmu.edu.cn (S. Fu).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.meegid.2016.04.016
1567-1348/

2002). The *MSH6* encodes protein hetero-dimerization and functions as a bidirectional molecular switch to repair DNA replication errors (Umar et al., 1998).

In fact, *MSH2*, *MSH6* and other MMR genes are associated with an increased risk of cancers such as prostate cancer (Chen et al., 2003), pancreatic cancer (Liu et al., 2014), gastric cancer (Yamamoto et al., 1999), lung cancer (Jung et al., 2006; Slovakova et al., 2015), non-small cell lung cancer (Hsu et al., 2007), breast cancer (Feng and Radivoyevitch, 2009), ovarian cancer (Quaye et al., 2009), endometrial cancer (EC) (Beiner et al., 2006; Imai and Yamamoto, 2008; Lacey et al., 2011; Poplawski et al., 2015), and hereditary nonpolyposis colorectal cancer (HNPCC) (Lynch and Lynch, 2000; Papadopoulos et al., 1994; Park et al., 1999; Peltomaki and Vasen, 1997; Watson et al., 1994). However, there are no functional and association studies that have systematically assessed the association between the MMR genes, HIV-1 infection and AIDS progression. To further investigate the role of the polymorphisms in the *MSH2* and the *MSH6* genes in the development of HIV-1 infection and AIDS, we performed an association study of 14 single nucleotide polymorphisms (SNPs) in the *MSH2* and the *MSH6* genes in 966 northern Chinese individuals. Participants were genotyped to investigate whether the two genes polymorphisms were associated with the susceptibility to HIV-1 infection and the progression of AIDS.

## 2. Material and methods

### 2.1. Subjects

In this study, 479 HIV-1-infected male individuals who have sex with men (MSM) were recruited from Heilongjiang Center for Disease Control and Prevention (CDC). The age of the HIV-1-seropositive individuals ranged from 16 to 75 years old (mean age ± SD, 35.3 ± 11.55) and the average $CD4^+$ T-lymphocyte count at that time point was 335 cells/μl (range, 3–1038 cells/μl). We categorized these patients as Category 1 (T-lymphocytes >200 cells/μl) or Category 2 (T-lymphocytes <200 cells/μl) by the $CD4^+$ T-lymphocyte count, and as Category A (Clinical phase I + II + III) or Category B (Clinical phase IV) by the clinical stage.

The age and sex frequency-matched unrelated healthy controls (n = 487) were recruited from individuals who were HIV-1-seronegative individuals diagnosed through a comprehensive medical test at the clinical laboratory of the second affiliated hospital of Harbin Medical University. The parents of each participant were individuals of the same nationality (for at least three generations) in non-sanguineous marriages. The age of the uninfected controls ranged from 16 to 75 years (mean age ± SD, 35.3 ± 11.59). All participants provided informed consent approved by the appropriate local authority.

### 2.2. SNPs selection and genotyping

We included candidate SNPs based on the published literature and chose tagging SNPs based on linkage disequilibrium (LD) in the HapMap (HapMap Data Rel 24/phase II Nov08, on NCBI B36 assembly, dbSNP b126). The tagging SNPs were screened for minor allele frequency (MAF) ≥10% among Han Chinese in Beijing, China (CHB), and no representation by other tagging SNPs at a LD of $r^2 \geq 0.80$. We selected nine SNPs in *MSH2* including rs13019654, rs4608577, rs4952887, rs6726691, rs10191478, rs12999145, rs1981929, rs2042649, and rs2303428 and five SNPs in *MSH6* gene including rs2348244, rs3136245, rs3136329, rs2072447, and rs7562048. Of the 14 SNPs, rs6726691, rs12999145, rs3136245, rs3136329, and rs7562048 were selected from the literature and the others were tagging SNPs. Peripheral blood of all participants was collected and genomic DNA was extracted using the QIAamp blood kit (Qiagen, Germany). The 14 SNPs were genotyped using a custom-design 48-Plex SNPscan™ Kit (Genesky Bio-technologies Inc., Shanghai, China). This kit was developed according to patented SNP genotyping technology by Genesky Biotechnologies Inc., and was based on a double ligation and multiplex fluorescence PCR. For quality control, a 5% random sample of cases and controls was genotyped twice to verify the genotyping accuracy, the reproducibility was 100%.

### 2.3. Statistical analysis

After the genotypes of all participants were determined, the genotype and allele frequency were calculated by directly counting the number. The Chi-square test was used for examining the deviation from Hardy–Weinberg equilibrium (HWE) for all SNPs of the healthy group, the association between the allele, genotype frequency and susceptibility to HIV-1 infection, and the SNPs and the clinical features of cases (such as the $CD4^+$ T-lymphocyte count and clinical stage). The haploView 4.2 software was used to evaluate the LD and the frequency of the haplotypes between HIV-1-infected individuals and health controls. The relative risk associated with rare alleles was estimated as an odds ratio (OR) and 95% confidence interval (CI). SPSS 17.0 (SPSS, Chicago, IL, USA) was used for all statistical analyses. The differences with a *P* value less than 0.05 were considered statistically significant, which was further subjected to a Bonferroni correction for multiple testing. Statistical power was assessed using the Genetic Power Calculator (Purcell et al., 2003). Considering 0.06% prevalence of the disease, a risk allele frequency of 17%, and an additive genetic model, we had at least 95% power to detect an OR of 1.5 at the 0.05 level.

## 3. Results

The information of the 14 SNPs and the allele frequency in the *MSH2* and the *MSH6* genes are shown in Table 1. All tested SNPs did not deviate from the Hardy–Weinberg equilibrium in the healthy group ($P > 0.05$). There were no significant differences in the allele frequency of the 14 SNPs between cases and controls ($P > 0.05$). Similarly, we found no associations between genotype frequency of the 14 SNPs and HIV-1 infection under different genetic models ($P > 0.05$) (Supplementary Table S1).

In analyzing the associations between the 14 SNPs and clinical features of AIDS (Table 2), we found that there was a significant association between the rs7562048 and the $CD4^+$ T-lymphocyte count in patients with the AIDS ($P < 0.05$). The frequency of the rs7562048 G allele was significantly higher in the cases with Category 2(T-lymphocytes < 200 cells/μl) compared with those with Category 1 (T-lymphocytes > 200 cells/μl) ($P = 0.001$, OR = 1.811, 95% CI 1.255–2.614). The difference between Category 1 and Category 2 was still statistically significance after the Bonferroni correction. There were significant associations of the rs2348244 and rs3136245 with the clinical stage of AIDS. The frequencies of the rs2348244 C allele and the rs3136245 T allele were higher in the cases with Category B (Clinical phase IV) than those with Category A (Clinical phase I + II + III) ($P = 0.026$, OR = 1.591, 95% CI 1.056–2.398 and $P = 0.019$, OR = 1.749, 95% CI 1.096–2.791, respectively). However, the significance between Category A and Category B disappeared after the Bonferroni correction.

There was strong LD between the nine SNPs in the *MSH2* gene as well as the five SNPs in the *MSH6* gene. Nine haplotypes in block 1 and eight haplotypes in blocks 2 and 3 were identified and the frequencies of these haplotypes are listed in Table 3. The frequency of the haplotype TCCAAGCTC in cases was higher than that in healthy controls ($P = 0.0498$). However, there were no differences in the frequencies of the other haplotypes between HIV-1-infected individuals and healthy cohorts and there was no significant association with the susceptibility to HIV-1 infection ($P > 0.05$).

## 4. Discussion

To our knowledge, this is the first comprehensive study to systematically evaluate the association of the *MSH2* and the *MSH6* genes