Thymidylate synthase enhancer region polymorphism not related to susceptibility to acute lymphoblastic leukemia in the Kashmir population

F.H. Nazki¹, A. Masood¹, M.A. Banday², A. Bhat¹ and B.A. Ganai¹

¹Department of Biochemistry, University of Kashmir, Srinagar, India
²Department of Medical Oncology, SKIMS, Srinagar, India

Corresponding author: B.A. Ganai
E-mail: bbcganai@gmail.com

Received July 29, 2011
Accepted December 23, 2011
Published April 10, 2012
DOI http://dx.doi.org/10.4238/2012.April.10.6

ABSTRACT. Thymidylate synthase (TS) is a crucial enzyme in folate metabolism and plays a vital role in DNA synthesis and repair. The most common polymorphism in TS is a unique double (2R) or triple (3R) 28-bp tandem repeat sequence in the enhancer region of the TS gene (TSER). This genetic variation in TSER has been widely investigated and has been implicated as a risk factor for the development of various cancers, including acute lymphoblastic leukemia. It has also been found to influence sensitivity to anti-cancer drugs, such as methotrexate. We evaluated this polymorphism in acute lymphoblastic leukemia patients in the Kashmir population. In order to determine whether a double (2R2R) versus a triple (3R3R) 28-bp tandem repeat in the TSER modulates risk for acute lymphoblastic leukemia, 72 acute lymphoblastic leukemia cases and 144 age and gender matched, unrelated healthy individuals from the Kashmir region of India were evaluated for this polymorphism by PCR and direct sequencing. We found the frequency of the TS 2R allele to be 32.6 and 26.0%, in cases and controls, respectively. The TS 2R/2R genotype was found to be present in 15.27% of the cases
and 9.72% of the controls, the 2R/3R variant in 34.72% of the cases and 32.63% of the controls, and the 3R/3R genotype in 50.0% of the cases and 57.63% of the controls. There was a significant association between the TS 2R/2R genotype and gender of acute lymphoblastic leukemia patients with males harboring the 2R/2R genotype exhibiting a higher risk of developing acute lymphoblastic leukemia than females (P = 0.009). We concluded that the TSER polymorphism appears not to be a risk factor for susceptibility to acute lymphoblastic leukemia in the Kashmir population.

Key words: ALL; TSER; 28-bp repeat; Kashmir; Polymorphism; India

INTRODUCTION

Acute lymphoblastic leukemia (ALL) is the most common form of cancer in children, accounting for 30% of all the pediatric malignancies (Karathanasis et al., 2009). Although the clinical and pathological aspects of leukemia are well studied, little is known about the genes that affect the susceptibility to this disease (Skibola et al., 2002). The development of ALL has been proposed to arise through a combination of genetic predisposition and exposure to environmental factors (Krajinovic et al., 2004). Consistent with this paradigm, variants of genes involved in xenobiotic metabolism, DNA repair pathway and cell cycle checkpoint functions have been shown to influence the susceptibility to ALL (Pui et al., 2008).

Folate is an essential nutrient for cellular functioning because it provides one-carbon donors that are required for DNA synthesis, methylation and repair (De Jong et al., 2009). Methotrexate, an antifolic acid agent, has proven to be an effective chemotherapeutic drug for the treatment of lymphoid malignancies, indicating an association between the folate metabolism and the development of such malignancies (Hishida et al., 2003). Many enzymes are involved in the folate metabolism, among which, thymidylate synthase (TS) is a crucial enzyme and hence a good candidate for studying the effect of polymorphisms in the folate metabolism gene on the development of malignancies.

TS, encoded by the TS gene located on chromosome 18p11.32, plays a vital role in maintaining a balanced supply of deoxynucleotides required for DNA synthesis and repair (Danenberg, 1977) by catalyzing the conversion of dUMP to dTMP. The most common polymorphism in TS is a unique double (2R) or triple (3R) 28-bp tandem repeat sequence in the 5' untranslated region (5'-UTR) of the TS gene also called TS enhancer region (TSER), immediately upstream from the initiation site, which influences protein expression in cancer cells. The presence of a triple versus double 28-bp repeat in the enhancer region has been associated with an increased TS expression both in in vivo and in vitro studies (Horie et al., 1995, Horie and Takeishi, 1997; Kawakami et al., 1999; Pullarkat et al., 2001). This increased expression may, in turn, increase the conversion of dUMP to dTMP, thereby, decreasing uracil levels and the consequent erroneous incorporation of uracil into DNA of rapidly dividing hematopoietic stem cells and could work protectively against the development of ALL (Skibola et al., 2002). The TS 28-bp repeat polymorphism has been shown to modulate the risk of ALL in various populations but the results obtained are controversial and require further investigation to be confirmed and clarified (Skibola et al., 2002; Lauten et al., 2003;
The aim of the present study was to determine the frequency of the TSER polymorphism and its possible association with the susceptibility to ALL in Kashmiri subjects (India). The frequency of TSER genotypes with respect to various clinico-pathological characteristics was also evaluated in this study. Kashmir is located in a small valley at a high altitude, with mostly consanguineous marriages resulting in preservation of a genetic pool. Owing to this factor and its geographical location Kashmiris show wide genetic diversity from the rest of India. Therefore, an evaluation of the TS polymorphism in the Kashmir population is required to determine the role of this gene in the susceptibility to ALL.

MATERIAL AND METHODS

Subjects

A case-control study was performed involving 72 patients with ALL and a control group composed of 144 individuals without leukemia. The ALL samples were obtained from the Department of Medical Oncology - SKIMS, Srinagar, collected sequentially between March 2009 and January 2011 from patients at the time of diagnosis. The diagnosis of ALL was based on French-American-British (FAB) or immunophenotypic criteria. The main characteristics of the patients are shown in Table 1. All cases of childhood and adult ALL were grouped because we did not find any statistical difference in the incidence of the TSER polymorphism between patients aged 16 years or younger (childhood ALL) and those older than 16 (adult ALL).

<table>
<thead>
<tr>
<th>Variant</th>
<th>ALL (N = 72)</th>
<th>Control group (N = 144)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤16</td>
<td>43; 10 (3-16)</td>
<td>86; 12 (3-16)</td>
<td></td>
</tr>
<tr>
<td>&gt;16</td>
<td>29; 22 (17-47)</td>
<td>58; 25 (17-45)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>44 (61.1)</td>
<td>82 (56.94)</td>
<td>0.55</td>
</tr>
<tr>
<td>Female</td>
<td>28 (38.8)</td>
<td>62 (43.05)</td>
<td></td>
</tr>
<tr>
<td>Dwelling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>61 (84.7)</td>
<td>121 (84.02)</td>
<td>0.89</td>
</tr>
<tr>
<td>Urban</td>
<td>11 (15.2)</td>
<td>23 (15.97)</td>
<td></td>
</tr>
<tr>
<td>French-American-British type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>36 (50.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>19 (26.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biphenotypic</td>
<td>3 (4.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not classified</td>
<td>14 (19.4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data are reported as number with percent in parentheses, except for age [number; median (range)]. ALL = acute lymphoblastic leukemia. P value using the chi-square test.

DNA isolation

Three milliliters of venous blood was drawn into a sterile tube containing ethylenediamine tetraacetic acid (EDTA) and stored at -20°C until the isolation of genomic DNA. DNA was isolated by a standard phenol-chloroform extraction method (Blin and Stafford, 1976).
**PCR analysis for TS 2R→3R genotyping**

The PCR primer conditions were the same as previously described (Zhang et al., 2004). PCR was performed in a 25-µL total volume reaction mixture containing 100 ng template DNA and 200 nM each of the forward (5’-GTGGCTCCTGCTTTCCCCC-3’) and reverse (5’-GGCTCCGAGCCGCACAGGCATGGCGCGG-3’) primers were used. PCR cycling parameters were a 5-min denaturation cycle at 94°C and 35 cycles of the following: 94°C for 30 s, 63°C for 30 s, and 72°C for 30 s, and a final extension at 72°C for 5 min. The amplified products were visualized on a 3% agarose gel using ethidium bromide.

**Direct sequencing**

DNA sequencing analysis was carried out to confirm the results of TS genotyping at the 5’-UTR tandem repeat locus in a subset of 12 representative samples. PCR amplicons were recovered from agarose gel, followed by purification with a Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Irvine, CA, USA) and then used for direct DNA sequencing. DNA sequencing was carried out at MACROGEN Inc. (Seoul, Korea). To rule out the possibility of sequencing artifacts by PCR, products from at least two different PCRs were sequenced using forward and reverse primers.

**Statistical analysis**

The chi-square test was used for the comparison of the allele and genotype frequency between the cases and controls. The distribution of the genotype frequencies in both groups did not deviate from the Hardy-Weinberg equilibrium. The odds ratios (OR) were calculated as estimates of relative risk for disease and 95% confidence intervals (CI) were calculated for all observed allele frequencies. A P < 0.05 was considered to be statistically significant. The SPSS statistical software package (ver. 11.5, SPSS, Chicago, IL, USA) was used for the statistical analysis.

**RESULTS**

Seventy-two histopathologically confirmed ALL cases and 144 matched controls belonging to the Kashmir group were analyzed for the TSER polymorphism. The homozygotes for the double repeat (2R/2R) produced a single band corresponding to 215 bp. Heterozygotes (2R/3R) produced two fragments corresponding to 215 and 243 bp, and homozygotes for the triple repeat (3R/3R) produced a single 243-bp fragment (Figures 1 and 2). Sequencing analyses showed that these fragments contained the same sequence except for the last 28-bp repeat of both 2R and 3R genotypes, which contains a 2-bp insertion (CC, site). A single G→C (SNP) at the 12th nucleotide of the second repeat of the 3R allele was also observed in some of the samples sequenced (Figures 3 and 4).
Figure 1. TSER genotyping by PCR amplification followed by separation on 3% agarose gel as described in text. 
*Lane M = 50-bp ladder; lanes 2-6 = PCR products of the TSER genotypes (3R/3R, 2R/3R, 2R/3R, 3R/3R, 3R/3R, 3R/2R); lane 7 = negative control.*

Figure 2. TSER genotyping by PCR amplification followed by separation on 3% agarose gel as described in text. 
*Lane M = DNA marker; lanes 2-7 = PCR products of the TSER genotypes (3R/3R, 3R/3R, 3R/3R, 3R/3R, 3R/2R, 3R/2R, 2R/2R).*
Figure 3. Partial sequence electropherogram of DNA sample (+strand) of normal control homozygous for the 3R repeat. The tandem repeats are demarcated by arrows and represented as 1, 2 and 3. The SNP at the 12th nucleotide position of the second repeat is represented by the black arrow. The insertion at the end of the last repeat is represented by CC.
Figure 4. Partial sequence electropherogram of DNA sample (+strand) of ALL patient homozygous for the 2R repeat. The tandem repeats are demarcated by arrows and represented as 1 and 2.

TS polymorphism

The TS triple tandem repeat (3R) allele frequency was found to be 73.95% in the controls and 67.36% in the cases. The difference in frequency was found to be statistically insignificant with a P value of 0.1510 (P > 0.05). The TS 2R/2R genotype was found to be present in 15.27% of the cases and 9.72% of the controls, the 2R/3R variant in 34.72% of the cases and 32.63% of the controls, and the 3R/3R genotype in 50.0% of the cases and 57.63% of the controls (Table 2).
It was observed that although the proportion of patients who were homozygous for the TS tandem repeat (3R/3R) was lower in cases than in controls, the difference was not statistically significant when using the 2R/2R genotype as a reference (OR = 0.552; 95%CI = 0.2287-1.3325; P = 0.1826). Similarly, it was observed that the frequency of the heterozygous genotype (2R/3R) when compared with the 2R/2R genotype was not much different between the cases and controls (OR = 0.677; 95%CI = 0.268-1.7105; P = 0.4081).

### TS genotypes and clinico-pathological variables

Table 3 shows the genotype frequencies of TSER polymorphism relating to age, gender, dwelling, and FAB type of patients. The TSER genotype was only statistically significantly when associated with the gender of the patients (P = 0.009); age, dwelling and FAB type showed no association with the TSER polymorphism.

<table>
<thead>
<tr>
<th>Variables (N = 72)</th>
<th>TS polymorphism</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2R/2R</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
</tr>
<tr>
<td>≤16</td>
<td>11 (15.27)</td>
</tr>
<tr>
<td>&gt;16</td>
<td>25 (34.72)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>36 (50.0)</td>
</tr>
<tr>
<td>Females</td>
<td>47 (32.64)</td>
</tr>
<tr>
<td>Dwelling</td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>97 (67.36)</td>
</tr>
<tr>
<td>Urban</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>French-American-British type</td>
<td>6 (10.34)</td>
</tr>
<tr>
<td>L2 + biphenotypic</td>
<td>6+0 (10.34)</td>
</tr>
</tbody>
</table>

Data are reported as number with percent in parentheses. OR = odds ratio; CI = confidence interval. P value using the chi-square test.

**DISCUSSION**

ALL is the most common cancer in childhood with an event-free survival rate of near to 80% in the developed countries (Pui et al., 2008). Folate plays an important role in DNA synthesis and repair. Folate deficiency results in the large scale misincorporation of uracil into...
DNA and chromosome breaks. This in turn induces chromosome damage, fragile site formation, micronucleus formation, and elevated uracil levels in the DNA of the bone marrow cells (Hori et al., 1984). Taking this information into consideration it seems plausible that functional alterations due to polymorphisms in the genes involved in folate metabolism may be associated with the development of cancers. Consistent with this hypothesis it has been shown that polymorphisms in the folate-related genes have been associated with the risk of cancers (Kim, 1999). TS is a crucial enzyme in this folate metabolism. It catalyzes the conversion of dUMP to dTMP using methylene THF as the carbon donor for the reaction and maintains the balanced supply of four precursor nucleotides for DNA synthesis and repair. Alteration in TS gene expression could affect this balanced supply predominantly in rapidly dividing cells such as hematopoietic stem cells. TS comprises a unique 28-bp repeat sequence in its 5'-UTR, which has been found to be polymorphic containing either two (2R) or three (3R) 28-bp repeats (Hishida et al., 2003). The less predominant 2R allele has been shown to result in 2.6 times lower gene expression than the 3R allele in in vitro studies (Horie et al., 1995) and 3.6-fold lower mRNA expression levels in tumor tissue of TS 2R/2R individuals than in tissue from individuals homozygous for the 3R condition (Pullarkat et al., 2001). These tandem repeat sequences in TS are presumed to regulate gene expression by forming secondary structures in the 5'-UTR of mRNA (Kawakami et al., 1999).

In the present study, the genetic association of this 5'-UTR tandem repeat polymorphism in the TS gene with the development of ALL in the Kashmir population was investigated. The repeat polymorphism in the TS gene was evaluated in 72 ALL cases and 144 (age, gender and region matched, non-malignant) controls by PCR analysis of DNA obtained from the blood of the subjects, followed by direct sequencing of DNA. We observed that the TS triple tandem repeat (3R) allele frequency was 73.95% in the controls and 67.36% in the cases. The difference in frequency was found to be statistically insignificant with a P value of 0.1510 (P > 0.05). The TS 2R/2R genotype was found to be present in 15.27% of the cases and 9.72% of the controls, the 2R/3R variant in 34.72% of the cases and 32.63% of the controls, and the 3R/3R genotype in 50.0% of the cases and 57.63% of the controls.

Thus, our study suggests that there is no association between TS tandem repeat polymorphism and the development of ALL in the Kashmir population. Similarly, on evaluating the distribution of TSER genotypes with respect to various clinico-pathological characteristics we found that the TS 2R/2R genotype was only statistically significantly associated with the gender of the subjects, with males harboring the TS polymorphism showing a higher risk than females of developing ALL (P = 0.009). The statistical results obtained for patients with ALL are based mainly on subjects older than 10 years (51/72; 71%). For this age group, a clear predominance of ALL among males has been previously reported (Cartwright et al., 2002). In this regard, our finding of the increased risk of ALL in males conferred by the TS 2R/2R genotype provides a genetic basis for the higher incidence of ALL reported in males. Since there is no current environmental hypothesis to explain the higher incidence of acute lymphoblastic leukemia in males, it is hypothesized that females are genetically better protected than males against the environmental and toxic agents that are involved in the development of acute lymphoblastic leukemia.

The results from several studies have shown an inconsistent association between genes involved in folate metabolism and the risk of ALL. It has been shown that the ethnicity of an individual plays a role in the various distributions of polymorphisms of the genes...
involved in folate metabolism (Lima et al., 2008; Giovannetti et al., 2008). The prevalence of TS 3R/3R has been reported to be 33.1% in Caucasian children with ALL, while in Indonesian samples, its frequency reached 76.1% in ALL children (Giovannetti et al., 2008). Furthermore, the frequency of TS genotypes in the Indonesian ALL patients and controls did not show any significant difference (Giovannetti et al., 2008). In the study of Skibola et al. (2002), individuals with TS 2R/3R were shown to exhibit a 2.8-fold reduction and those with TS 3R/3R exhibited a 4.0-fold reduction in ALL risk. In contrast, in a western European pediatric series of ALL patients, the TS 2R variant was reported to have a protective effect (De Jong et al., 2009).

A study, involving 40 ALL Canadian patients and 40 controls, reported that the frequency of TS genotypes was not significantly different (Lauten et al., 2003). Also, in a large number of ALL patients from the United Kingdom Childhood Cancer study, the frequency of TS 28-bp repeat was not significantly different compared to controls (Lightfoot et al., 2010). Similarly, in another study involving 73 ALL patients and 128 controls from a western Iranian population, it was observed that the frequency of the 2R allele was not significantly higher in ALL patients as compared to controls (Rahimi et al., 2010). However, an interaction between the TS 28-bp repeat polymorphism and folate intake has been reported with a decreased risk of ALL in the 3R/3R genotype in combination with high folate intake (Ulrich et al., 2002).

Petra et al. (2007) in Slovenia, reported that TS polymorphisms were not significantly associated with ALL risk.

The reasons for contrary results obtained from several studies remain ambiguous and might be attributed to differences in ethnic backgrounds and the selection of the population studied, differences in sample sizes, and gene-environment interactions, such as diet, exposure to chemicals, or nutritional intake of folate and related vitamins.

In addition, gene-gene interactions and the presence of an additional G→C SNP within the second repeat of the triple tandem have been shown to influence the transcriptional activity of the TS gene (Mandola et al., 2003). In our study, sequencing results of some subjects did show the presence of such an SNP in the second repeat of triple tandem, which may influence the transcriptional activity of the gene and hence explain the reason for contrary results obtained. However, this needs further evaluation and investigation to establish the role of SNP in the development of ALL in the Kashmir population.

Furthermore, inherited biases accompanied with hospital-based case-control studies may also be attributed to spurious findings and false-positive results. Due to limited studies reporting the influence of gene polymorphism involved in folate metabolism on ALL, further investigations are required to find a clear relationship of these genes to susceptibility to ALL (Koppen et al., 2010).

In summary, our study has shown the frequency of the TS gene involved in folate metabolism with respect to ALL patients within a homogenous ethnic group. It seems that the TS 28-bp repeat polymorphism is not a risk factor for the susceptibility to ALL in the Kashmir population. However, additional analysis with a larger sample size is needed to clarify the real contribution of this gene in the susceptibility to ALL in different world populations.

ACKNOWLEDGMENTS

We are grateful to Dr. Arshid Butt, Muzamil Makhdoomi and Beenish Iqbal of the Department of Biochemistry, Kashmir University, for helpful suggestions. We would also like
to thank Dr. Aga Syed Sameer, Senior Research Fellow, Sher-i-Kashmir Institute of Medical Sciences, Srinagar, for his immense help in making this manuscript ready for publication by extensively editing and suggesting necessary changes. Research supported by grants from the Centre for Scientific and Industrial Research (CSIR), India, under the Junior Research Fellowship (JRF) scheme.

REFERENCES


