Mutation Screening and Association Study of the Folylpolyglutamate Synthetase (FPGS) Gene with Susceptibility to Childhood Acute Lymphoblastic Leukemia

Duangjai Piwkham¹², Teerapong Siriboonpiputtana¹, Joke Beuten², Samart Pakakasama³, Jonathan AL Gelfond⁴, Karan Paisooksantivatana¹, Gail E Tomlinson², Budsaba Rerkamnuaychoke¹*  

Abstract

**Background:** Folylpolyglutamate synthetase (FPGS), an important enzyme in the folate metabolic pathway, plays a central role in intracellular accumulation of folate and antifolate in several mammalian cell types. Loss of FPGS activity results in decreased cellular levels of antifolates and consequently to polyglutamatable antifolates in acute lymphoblastic leukemia (ALL). **Materials and Methods:** During May 1997 and December 2003, 134 children diagnosed with ALL were recruited from one hospital in Thailand. We performed a mutation analysis in the coding regions of the FPGS gene and the association between single nucleotide polymorphisms (SNPs) within FPGS in a case-control sample of childhood ALL patients. Mutation screening was conducted by polymerase chain reaction-single strand conformation polymorphism (PCR-SSCP) and subsequently with direct sequencing (n=72). Association analysis between common FPGS variants and ALL risk was done in 98 childhood ALL cases and 95 healthy volunteers recruited as controls. **Results:** Seven SNPs in the FPGS coding region were identified by mutation analysis, 3 of which (IVS13+55C>T, g.1297T>G, and g.1508C>T) were recognized as novel SNPs. Association analysis revealed 3 of 6 SNPs to confer significant increase in ALL risk these being rs7039798 (p=0.014, OR=2.14), rs1544105 (p=0.010, OR=2.24), and rs10106 (p=0.026, OR=1.99). **Conclusions:** These findings suggested that common genetic polymorphisms in the FPGS coding region including rs7039789, rs1544105, and rs10106 are significantly associated with increased ALL risk in Thai children.  

Keywords: Folylpolyglutamate synthetase - acute lymphoblastic leukemia - single nucleotide polymorphism

Asian Pac J Cancer Prev, 16 (11), 4727-4732

Introduction

Acute lymphoblastic leukemia (ALL) is a heterogeneous disease of hematological malignancies and recognized as the most common malignancy in children. Several risk factors have been shown to affect the risk of childhood acute lymphoblastic leukemia (Buffler et al., 2005; Emerenciano et al., 2007; Inaba and Mulligha, 2013) and a complex interplay between inherited genetic background and specific environmental exposures are most likely to determine ALL susceptibility (de Jonge et al., 2009; Yang et al., 2010). Genetic polymorphisms in genes involving in folate metabolic pathway are putative candidate modulators of ALL susceptibility (Thirumaran et al., 2005; Karathanasis and Kalmanti, 2009; Koppen and Kaspers, 2010; Yang et al., 2011). Folate metabolism provides one-carbon donors for the synthesis of de novo purines and pyrimidines necessary for RNA and DNA synthesis, remethylation of homocysteine, and methylation reactions (Wagner, 2001; de Beaumais, 2012). Low folate concentration may influence the risk of cancer through mechanisms of uracil misincorporation into DNA, possibly leading to double-strand breaks and chromosomal damage (Duthie, 1999; Duthie et al., 2004). Several studies have demonstrated the effects of polymorphism in folate metabolism genes with ALL risk. Koppen 2010, revealed that two common polymorphisms within the methylenetetrahydrofolate reductase (MTHFR) gene, MTHFR 677C>T (rs1801133) and MTHFR 1298A>C (rs1801131) are associated with a decreased susceptibility of childhood ALL (Koppen et al., 2010). Recently, Li 2014 using a case control study in 98 childhoods ALL and 93 age- and sex-matched non-ALL controls revealed that MTHFR 1298C alleles (AC or CC) were significantly increased a risk of ALL compared to the AA genotype (1.1 times). In addition,
frequencies of MTHFR 1298AA, 1289A>C, and 1298CC were significantly different in those two groups. However, no significant different in frequencies of MTHFR 677CC, 677C>T, and 677TT genotypes were observed in this report (Li et al., 2014). Moreover, MTHFR 1298A>C genotype and the 677CC/1298A>C haplotype were significantly associated with reduced the risk of acute myeloid leukemia (AML) compared with AA genotype and 677CC/1298AA haplotype. Furthermore, MTHFR 677TT was associated with an increased the risk of AML in non-smokers and non-drinkers (Huang et al., 2015). Case-control studies by meta-analysis were done in 6,371 cases and 10,850 controls by Jiang et al. The result supports that the MTHFR C677T genotype is associated with ALL risk in Caucasians (Jiang et al., 2013). On the other hand, no significant association between polymorphisms in MTHFR and MDR1 genes was reported in 68 childhood ALL (Kreile et al., 2014). Other enzymes involving folate metabolic pathway such as SHMT1 (cytosolic serine hydroxymethyltransferase) and thymidylate synthase (Skibola et al., 2012) have been reported to associate with the risk of ALL. SHMT1 1420T and TS 3R allele have been reported to associate with lower risk of ALL (de Jonge et al., 2009; Skibola et al., 2012). Furthermore, polymorphisms in SLC19A1 including 80A/A and 80G>A increased susceptibility to develop childhood ALL and associated with poor prognosis (de Jonge et al., 2009).

Folypolyglutamate synthetase (FPGS), is an essential enzyme for the survival of proliferating mammalian cells and critical for the conversion of chemotherapeutic agent, methotrexate (MTX), into MTX polyglutamate (Gómez-Gómez et al., 2014). The human FPGS gene, located on chromosome 9, consists of 15 exons spanning 11.2 kb of genomic DNA (Taylor et al., 1995; Chen et al., 1996). FPGS is present in several normal tissues, including bone marrow stem cells, liver, kidney, as well as in tumors (Moran and Colman, 1984; Barredo and Moran, 1992). The sensitivity of human tumors to antifolate chemotherapy was shown to be related to the level of expression of FPGS. Increased expression of FPGS activity leads to increased sensitivity to antifolate drugs while decreased FPGS activity is associated with resistance to many antifolates (Chen et al., 1996; Liani et al., 2003; Panetta et al., 2010). However, there is less clear whether genetic polymorphism within FPGS coding sequencing could contribute to the risk of leukemia development.

In this report, we aimed to investigate the association between the FPGS polymorphism and the susceptibility of ALL using a case-control group consisting of children diagnosed with ALL and children without malignancy matched by gender.

Materials and Methods

Subjects

One hundred and thirty four children diagnosed with ALL between May 1997 and December 2003 at Ramathibodi Hospital were included in this study. We used 39 cases for FPGS mutation screening. For association study, there were 52 males and 43 females with a median age of 6.2 years (range 5 months-14 years). The diagnosis of ALL was made by bone marrow morphology and immunophenotype. A total of 131 healthy volunteers comprising 50 males and 48 females were recruited as control in association study and 33 samples were used in FPGS mutation screening. Approval from Research Ethics Committee on Human Experimentation of Ramathibodi Hospital of Mahidol University was attained. Parental consent was verified for each child to participate in genetic studies. All samples underwent standard DNA preparation procedures and DNA was stored at 4°C prior to mutation screening and genotyping.

Polymerase chain reaction amplification

Based on the genomic reference sequence for FPGS [GenBank:NC_000009.11], 13 pairs of primers were designed using Perlprimer (http://perlprimer.sourceforge.net) to amplify all 15 exons including intron/exon boundaries, and 5' untranslated (5'-UTR) and 3'-UTR regions. Polymerase chain reactions (PCR) was performed in a total of 25 μl containing 0.5-1 μg genomic DNA, 0.4 pmol of each primer, 200 pmol dNTPs, and 1 unit AmpliTaq gold polymerase [Applied Biosystems, foster City, CA] (Daga et al., 1992). PCR condition was pre-denaturation at 94°C for 10 minutes, followed by 35 cycles of amplification cycles comprising denaturation at 94°C for 30 seconds, annealing at a temperature between 57 and 65 °C (Table 1) for 30 seconds, extension at 72°C for 1 minute and a final extension at 72°C for 10 minutes. PCR products were verified by 2% agarose gel electrophoresis.

Single strand conformation polymorphism (SSCP)

SSCP analysis was performed according to the method described by Bastos et al 2000 with slight modifications (Estela et al., 2001). Briefly, heat denatured PCR samples (1 μl) mixed with 2 μl SSCP gel loading dye were applied to a 12 % polyacrylamide gel (19 acrylamide: 1 bisacrylamide) and electrophoresis was carried out with a mini-PROTEAN II cell (Bio-Rad, Berkeley, CA) at a constant power of 120 V for 2 to 4 hours. The gel was stained with silver solution to verify the band pattern.

SNP selection and genotyping

Single nucleotide polymorphisms (SNPs) were selected from the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/) and SNPper (http://snpper.chip.org/) with the following criteria: a) SNPs with minor allele frequency (MAF)>0.05 that lead to an amino acid substitution or with potential functional effects, and (b) SNPs associated with leukemia, and in particular childhood ALL, as retrieved from literature, were included when available. Tagging SNPs within FPGS were identified using Haplovie [33] with the following criteria: (a) inclusion of preselected SNPs (see above), (b) MAF>0.05 in order to gain more statistical power, (c) r² threshold of 0.8 (d) a log of odds [34] threshold for multimarker testing of 3.0, (e) a minimum distance between tags of 60 bp, and (f) 2- and 3-marker haplotype tagging option. The search for tagging SNPs extended to a 10 kb region surrounding the gene. For rs7039798, rs7033913, rs1544105 and rs10106, genotyping was performed with the Golden Gate
Assay of the VeraCode technology using the BeadXpress reader System according to the manufacturer’s protocol (Illumina, San Diego, CA). SNP g.732C>T was genotyped using the TaqMan® allelic discrimination assay (Applied Biosystems, Foster City, USA) and for IVS13+55C>T, we used PCR-restriction fragment length polymorphism (PCR-RFLP) with corresponding restriction endonuclease AcI (New England Biolabs, Beverly, MA) for genotyping. After endonuclease digestion, the presence of specific PCR products was verified on a 2% agarose gel. Duplicate samples were included as quality controls.

**Statistical analysis**

Allele frequencies among the cases and controls were compared using Chi-square test ($\chi^2$) from Haploview. The odds ratio (OR) and its 95% confidence interval (Emerenciano et al., 2007) was estimated by logistic regression using R statistical software. All statistical tests were two-sided and significance was set at $P<0.05$. We used additive, dominant and recessive models in the test analysis. Haploview (http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploviev) was used to measure linkage disequilibrium and to define the block structure between the tagging SNPs (Barrett et al., 2005).

**Results**

**Mutation screening**

We performed PCR-SSCP and DNA re-sequencing of all 15 exons of FPGS including intron/exon boundaries, 5'UTR, and 3'UTR in 39 childhoods ALL and 33 controls. Seven SNPs were identified within FPGS gene as show in table 1. Interestingly, we could identify 3 novel SNPs including IVS9+19G>A, g.789C>T, and g.1508C>T which were not significantly different in allele frequency between cases and control (p > 0.05), as shown in table 2. The details of the identified SNPs including location, allele frequencies, and type of SNPs were summarized in table 1. In addition, five SNPs were recognized as coding SNPs which are located in exon 10 (g.732C>T and g.789C>T), exon 14 (g.1297T>G and exon 15 (g.1508G>T) and g.2006T>C), respectively. Two FPGS variants including g.1297T>G (exon 14) and g.1508C>T (exon 15) resulted in amino acid change including F433V and A503V, respectively. In contrast, two synonymous SNPs including g.732C>T and g.789C>T were observed in FPGS exon 10. Moreover, two SNPs identified in our mutation analysis were located in the intron boundaries (IVS9+19G>A and IVS13+55C>T) or in the 3'UTR region (rs10106). The allele frequency distribution showed that the majority of the variants were presented at low frequency. Interestingly, FPGS variants including IVS9+19G>A, g.789C>T, and g.1508C>T were found to be polymorphic in ALL cases only, whereas g.1297T>G was restrict observed in controls. To gather, we demonstrated the distribution of genetic variants (SNPs) within FPGS gene in childhoods ALL and normal children in Thailand. Two FPGS variants including F433V and A503V may associate to the risk of ALL in Thai patient especially the A503V which restrict observed in ALL patients.

**Association analysis between FPGS variants and ALL risk**

We performed association analysis on 6 FPGS variants which are comprised of our identified 2 SNPs (g.732C>T and IVS 13+55C>T) and 4 common FPGS tagging SNPs database (rs7039798, rs7033913, rs1544105, and rs10106). Those SNPs covered the FPGS gene (from the promoter to the 3'UTR region) and included the 10 kb flanking region of the gene. All tested SNPs were in Hardy Weinberg Equilibrium (HWE) and their allele frequencies were shown in table 2. Strong linkage disequilibrium was found between the 6 markers (pairwise $D'=1.00$ except between exon 10 and Intron 13 SNPs). All markers were part of one haplotype block in the controls whereas in cases a haplotype block was present encompassing rs7039798-rs7033913-rs1544105 (data not shown). Although no significant difference was observed in allele frequency distributions between cases and controls, single variant analysis revealed 3 SNPs were significantly associated with ALL risk (p < 0.05). Those were the AG genotype of rs7039798 (p=0.014, OR=2.14, 95%CI=1.16-3.94), the AG genotype and G allele of rs1544105 (p=0.010, OR=2.24, 95%CI=1.22-4.1 and p=0.038, OR=1.83, 95%CI=1.03-3.24, respectively), and the TC genotype of rs10106 (p=0.026, OR= 1.99, 95%CI=1.08-3.64), respectively (table 3). To summarize, we could identify FPGS variants (SNPs) that may associate to the risk of...
Methotrexate (MTX) is recognized as one of primary anticancer therapeutic agents for a treatment of childhood ALL. The better understanding in molecular underlining biology of ALL as well as molecular dissections of pathway involving folate metabolism have shaped the way to improve the treatment of ALL. Low FPGS activity results in decreased of antifolate in the cell and leading to drug resistance in leukemia (Stark et al., 2009). Here, we investigated the genetic polymorphisms of FPGS gene which is recognized as a critical component of folate metabolic pathways in Thai children with ALL as well as gender- and age-specific control samples. Furthermore, we performed the association study to determine the risk of ALL using our identified SNPs combined with the analysis of common variants SNPs within FPDS gene.

In this work, conventional PCR-SSCP were performed as a screening test to determine FPGS polymorphisms in each individual following with direct sequencing to confirm each FPGS variant. We could identify 7 SNPs within the coding region of FPGS gene (15 exons with intron/exon boundaries). Interestingly, we were able to identified 3 novel SNPs including IVS13+55C>T, g.732C>T, and g.1508C>T. These results suggested that polymorphism of FPGS including IVS13+55C>T, g.732C>T, and g.1508C>T may unique in Thai population. Furthermore, genotyping of 2 FPGS SNPs in this work combined with 4 tagging SNPs covering FPGS gene in 193 samples (95 cases and 98 controls), revealed that significant increasing of ALL risk was observed in FPGS rs7039798 (p=0.014, OR=2.14), rs1544105 (p=0.010, OR =2.24), and rs10106 (p=0.026, OR=1.99), respectively. To gain more confidence, comparative analysis with other studies in the genetic polymorphisms of FPGS is necessary.

There are little known about the role of FPGS variants among population in the establishment of cancer risk as well as in ALL. In this report, we demonstrated that IVS9+19G>A, g.789C>T, and g.1508C>T were unique observed in ALL cases, whereas g.1297T>G was restrict found in controls. Interestingly, FPGS 1508C>T (exon 15) resulted in amino acid substitution (A503V) which is specifically observed in case group may play a role to determine the genetics risk of ALL in our tested population. Recent studies in FPGS polymorphisms in ALL revealed that. FPGS rs1544105C>T polymorphism might influence FPGS expression and affect treatment outcome in ALL patients (Liu et al., 2013). In addition, FPGS A22G is significantly increased the risk of ALL and may affect the overall survival of ALL patients (Gómez-Gómez et al., 2014). To gather, several genetic polymorphisms in FPGS genes were observed to be associated with the risk of ALL and these variants were largely depended on the different in genetic background of tested populations.

For the perspective in molecular mechanisms of FPGS in the establishment of ALL, Raz, 2014 demonstrated that mutations of FPGS exon 12 altered the transcription levels of FPGS in several cell lines and these exon could function as the transcriptional regulatory of FPGS via several mechanisms such as transcriptional binding domain and hypermethylation site (Raz et al., 2014). However, further functional studies of each FPGS variant

### Discussion

Methotrexate (MTX) is recognized as one of primary anticancer therapeutic agents for a treatment of childhood ALL. The better understanding in molecular underlining biology of ALL as well as molecular dissections of pathway involving folate metabolism have shaped the way to improve the treatment of ALL. Low FPGS activity results in decreased of antifolate in the cell and leading to drug resistance in leukemia (Stark et al., 2009). Here, we investigated the genetic polymorphisms of FPGS gene which is recognized as a critical component of folate metabolic pathways in Thai children with ALL as well as gender- and age-specific control samples. Furthermore, we performed the association study to determine the risk of ALL using our identified SNPs combined with the analysis of common variants SNPs within FPDS gene.

In this work, conventional PCR-SSCP were performed as a screening test to determine FPGS polymorphisms in each individual following with direct sequencing to confirm each FPGS variant. We could identify 7 SNPs within the coding region of FPGS gene (15 exons with intron/exon boundaries). Interestingly, we were able to identified 3 novel SNPs including IVS13+55C>T, g.1297T>G and g.1508C>T. These results suggested that polymorphism of FPGS including IVS13+55C>T, g.1297T>G and g.1508C>T may unique in Thai population. Furthermore, genotyping of 2 FPGS SNPs in this work combined with 4 tagging SNPs covering FPGS gene in 193 samples (95 cases and 98 controls), revealed that significant increasing of ALL risk was observed in FPGS rs7039798 (p=0.014, OR=2.14), rs1544105 (p=0.010, OR =2.24), and rs10106 (p=0.026, OR=1.99), respectively. To gain more confidence, comparative analysis with other studies in the genetic polymorphisms of FPGS is necessary.

There are little known about the role of FPGS variants among population in the establishment of cancer risk as well as in ALL. In this report, we demonstrated that IVS9+19G>A, g.789C>T, and g.1508C>T were unique observed in ALL cases, whereas g.1297T>G was restrict found in controls. Interestingly, FPGS 1508C>T (exon 15) resulted in amino acid substitution (A503V) which is specifically observed in case group may play a role to determine the genetics risk of ALL in our tested population. Recent studies in FPGS polymorphisms in ALL revealed that. FPGS rs1544105C>T polymorphism might influence FPGS expression and affect treatment outcome in ALL patients (Liu et al., 2013). In addition, FPGS A22G is significantly increased the risk of ALL and may affect the overall survival of ALL patients (Gómez-Gómez et al., 2014). To gather, several genetic polymorphisms in FPGS genes were observed to be associated with the risk of ALL and these variants were largely depended on the different in genetic background of tested populations.

For the perspective in molecular mechanisms of FPGS in the establishment of ALL, Raz, 2014 demonstrated that mutations of FPGS exon 12 altered the transcription levels of FPGS in several cell lines and these exon could function as the transcriptional regulatory of FPGS via several mechanisms such as transcriptional binding domain and hypermethylation site (Raz et al., 2014). However, further functional studies of each FPGS variant
in the establishment of ALL and other types of cancer are still required.

Regarding SNPs selection and validation, 3 selected tagging SNPs were chosen according to HapMap data of the Caucasian in which the ethnic-specific LD patterns are observed. Thus, we might loss some tagged SNPs commonly observed in Thais. In addition, those selected SNPs in the study may not fully represent all tagged variants in this ethnic group. Nevertheless, selection of SNPs in the present study was enriched with potentially functional important SNPs. Therefore, those high priority SNPs were considered to include into this work. However, the power of this study was limited by several factors including sample size, the baseline incidence of the disease, and the unknown OR of a genetic risk factor. Replication of our findings in independent samples was warranted to elucidate the role of these variants in ALL susceptibility and defined their importance in the ethnic specific differences in ALL risk. Unfortunately, we did not assess the folate level or dietary intake of folate in our samples. Moreover, gene-environment interactions such as dietary records during pregnancy and their relation to the above mentioned polymorphisms are needed for the future cohort study with a larger number of cases.

In conclusion, our results suggested that common genetic variants in the FPGS genes were associated with ALL risk in Thai children. To our knowledge, this is the first report of an association between ALL risk and FPGS variants in the country. Our findings might help to improve the ability in assessing an individual risk for ALL as well as to elucidate leukemia risks at the population level.

Acknowledgements

This work was supported by the Thailand Research Fund (TRF) Grant # RMU5180024 (to BR), the Texas Higher Education Coordinating Board Grant #010019-0105-2001, the Children’s Cancer Fund Dallas, TX and the Greehey Family Foundation (to GET), and the Commission on Higher Education Scholarship, Thailand (to DP).

References

Raz S, Stark M and Assaraf YG (2014). Binding of a Smad4/ Ets-1 complex to a novel intragenic regulatory element in exon12 of FPGS underlies decreased gene expression and


