Identification of DLG5 and SLC22A5 gene polymorphisms in Malaysian patients with Crohn’s disease

Kek Heng CHUA,* Lay Hoong LIAN,* Boon Pin KEE,* Chooi Mei THUM,* Way Seah LEE,† Ida HILMI‡ & Khean Lee GOH‡

*Department of Molecular Medicine, †Department of Pediatrics, and ‡Division of Gastroenterology and Hepatology, Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

OBJECTIVE: The aim of this study was to investigate the association of DLG5 and SLC22A5 gene polymorphisms with the onset of Crohn’s disease (CD) in a Malaysian cohort.

METHODS: Genomic DNA of 80 CD patients and 100 healthy unrelated control individuals was extracted and analyzed via polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) on DLG5 (4136 C/A), DLG5_e26 and SLC22A5 (-207 G/C) genetic polymorphisms. Data obtained from the study were then subjected to statistical analysis to test for risk association.

RESULTS: Significant associations of both DLG5 polymorphisms with the development of CD in the Malaysian patients were observed in this study. The homozygous C genotype of the DLG5 polymorphism was significantly related to CD patients (P = 0.0023, OR = 2.5320), while the homozygous A was significant in control individuals (P = 0.0224, OR = 0.4480). In DLG5_e26 polymorphisms, we found a significant distribution of the homozygous insA genotype in CD patients (P = 0.0006, OR = 2.8916), whereas the heterozygous insA/delA genotype was significant in controls (P = 0.0007, OR = 0.3487). We hypothesized that there might be a complex interaction of both alleles, which conferred a protective effect against the onset of CD. However, we did not observe any significant correlation of SLC22A5 polymorphisms with this disease.

CONCLUSIONS: In our study, both polymorphisms in the DLG5 gene were found to be associated with CD patients in Malaysia. Therefore, these loci can be potentially used as susceptibility markers in the Malaysian population.

KEY WORDS: Crohn’s disease, DLG5, Malaysian, polymorphism, SLC22A5.

INTRODUCTION

Inflammatory bowel disease (IBD) is a medical condition that involves the inflammation of the small intestine and colon. There is a high prevalence rate of this condition in developing countries.1 Crohn’s disease (CD) is one of the major types of IBD, apart from ulcerative colitis. Patients with CD suffer from inflammation at multiple locations in the gastrointestinal tract, from mouth to anus.2 In addition,
CD also inflicts complications that involve other organs, such as eyes, joints, blood, skin and endocrine system. In clinics, patients often complain of diarrhea, rectal bleeding, pain in the lower right abdomen and weight loss. Unfortunately, there is yet to be a conclusive cure for CD. Medication and surgical treatments are given to patients for symptomatic relief and maintaining remission, and these can lead to a series of side effects such as nausea and skin rashes. CD is commonly observed in countries such as northern Europe, UK and North America. The prevalence rate of CD has been relatively low in developing countries in Asia. Recently, however, several studies had reported a rising trend of CD in Sri Lanka, Japan and Singapore. Scientists postulated that this phenomenon might result from an adoption of modern lifestyles from western counterparts. The overall prevalence of CD in Malaysia is 26 per 100 000, according to a study carried out in 2006. Interestingly, the prevalence rate differs among the three main ethnic groups that make up the multiracial society of Malaysia, that is Malay, Chinese and Indian. CD is more commonly observed in Indians than Chinese and Malays, with prevalence rates of 52.6, 26.9 and 9.3 per 100 000 individuals, respectively.

Despite the comprehensive researches that have been carried out by scientists worldwide, the exact etiology of CD remains unknown. In general, the onset of CD is believed to be triggered by a series of complex interactions of multiple factors. Both environmental factors and genetic composition are suggested to have a combined effect on an individual’s susceptibility to CD. Several potential genes have been observed to have direct association with the disease, namely CARD15 and IL23 genes. These genes play a role in the synthesis of proteins that are involved in immune responses. Variations in these regions, therefore, could lead to abnormal immune responses that subsequently result in chronic inflammation of the gastrointestinal tract. Recently, several other genes on chromosomes 5 and 10 have been identified as contributing to the risk of CD development. However, due to the complexity of genetic interaction in the disease, there is not a single gene known to be responsible for the onset of CD. Therefore, more extensive studies are being carried out to understand the role of various genes in the development and progression of the disease.

Recently, we have published data identifying single nucleotide polymorphism (SNP) 5 and JW1 mutations in the CARD15 gene in patients with CD. However, as with other large Asian studies, the common disease-predisposing mutations identified in the Caucasian population were not identified in our group of patients. In the present study, we investigated the distribution of variants of two SNPs present in the discs large homolog 5 (DLG5) and solute carrier 22A5 (SLC22A5) genes. The DLG5 gene is mapped to chromosome 10q23 and encodes for the membrane-associated guanylate kinase (MAGUK). The MAGUK contains four post-synaptic density protein, Drosophila disc large tumor suppressor, zonula occludens protein (PDZ) domains, a Src homology 3 (SH3) and guanylate kinase (GK) domains, as well as N-terminal and C-terminal domains. These domains are involved in the interaction of cells and proteins. The DLG5 gene consists of 32 exons of about 79 kb. There is evidence that DLG5 functions in maintaining the integrity and polarity of the cells and is also involved in the maintenance of the epithelial structure. Thus, any mutation in the DLG5 gene may hinder the normal function of the epithelial barrier in the gastrointestinal tract, making it more vulnerable to infection and disease.

An analysis of extended DLG5 haplotypes has identified four common haplotypes, that is A, B, C and D. In the present preliminary study, we focused on haplotype A tagging SNP (DLG5_e26) and haplotype C (4136 C/A). DLG5_e26 presents insertion or deletion of adenine (insA or delA) in exon 26, while DLG5 (4136 C/A) presents a C to A transversion in exon 23.

The SLC22A5 gene, on the other hand, is located on chromosome 5q31 and consists of 10 exons. The SLC22A5 gene encodes for the organic cation transporter 2 (OCTN2) and plays a main role as a transporter protein for organic cations, transporting carnitine (which is an essential cofactor of the metabolism of lipids). Thus, OCTN2 is important in the energy production of the cells. The presence of any genetic variations in the SLC22A5 gene could affect the expression levels of OCTN2 and subsequently lead to the obstruction of carnitine transportation. Without adequate amounts of carnitine, lipid metabolism in the cell cannot be carried out and thus, energy production ceases. This may then lead to characteristic signs and symptoms of CD. It has been reported that 28 genetic variants were found in the SLC22A5 gene: 2 are located in the 5’-untranslated region, 10 in the introns, 14 in the coding exons, one in the 3’-untranslated region and one in the promoter region. In the present study we examined the SLC22A5 (-207 G/C) polymorphism located in the promoter region.
MATERIALS AND METHODS

Patient and control sample collection

A total of 80 CD patients were randomly selected from the University Malaya Medical Centre (UMMC) located in Kuala Lumpur, Malaysia. Another 100 gender and race-matched unrelated healthy individuals were recruited as controls. This study was approved by the UMMC Ethics Review Board (approval number: 472.55) and informed consents were obtained from all participants. Genomic DNA was extracted from peripheral blood samples using a conventional phenol chloroform extraction method that has been applied previously.25,26 The quality and quantity of extracted DNA were accessed via spectrophotometry.

Polymerase chain reaction (PCR) amplification of DLG5 (4136 C/A), DLG5_e26 and SLC22A5 (-207 G/C)

Primer sequences used in this study to type all the three polymorphisms, i.e., DLG5 (4136 C/A): forward: 5′-AGCTCACACCTGGACCCTGCCGGTAC-3′ and reverse: 5′-TCACAGCAACGTCTGCTGACCTGGAGCTCCACTGC-3′; DLG5_e26: forward: 5′-CGACATCCTACGTTGGATGACACCTTACC-3′ and reverse: 5′-AGATAAGAGAGTTTCTTCCTACCCACCAGC-3′; SLC22A5 (-207 G/C): forward: 5′-GCGCCGCTCTGCTGAGTGC-3′ and reverse: 5′-AGGTAGGGCTGCGCAAGCTGACACC-3′, were synthesized according to a previous study.12 In silico PCR as described previously was also performed to ensure no unspecific binding to other loci prior to the actual PCR.27–29 Cycling parameters were optimized to the desired conditions. The PCR was carried out in a 20 μL reagent mixture containing 1 x Taq buffer with potassium chloride, 1 unit of Taq DNA polymerase (Fermentas, Glen Burnie, MD, USA), 2 mmol magnesium chloride, 0.5 mmol each of deoxynucleotide triphosphate mixture, forward and reverse primers and 50 ng genomic DNA. The cycling program was set according to our standard PCR conditions as described previously.30

Restricted fragment length polymorphisms (RFLP) and statistical analysis

Post-PCR amplicons were subjected to restriction enzyme (RE) digestion. The RE used for different gene polymorphism studies with the expected digested product sizes are shown in Table 1. The digested products were electrophoresed on a 3% (w/v) ethidium bromide-stained agarose gel for analysis. The genotype and allelic frequencies were calculated to correlate these polymorphisms to the onset of CD in the Malaysian population. χ², P value, odds ratio (OR) and 95% confidence interval (CI) were also calculated using SPSS 12.0 (SPSS Inc., Chicago, IL, USA) to access the risk of CD in individuals with particular traits. The value of P < 0.05 was defined as statistically significant.

RESULTS

DLG5 (4136 C/A) polymorphisms

Figure 1 shows the banding patterns for the polymorphisms in DLG5 (4136 C/A) after RE digestion. Three
genotypes were observed in the population, that is homozygous C, heterozygous C/A and homozygous A. The genotype and allelic frequencies of the DLG5 (4136 C/A) polymorphisms are shown in Table 2. We found significant differences in both genotypic and allelic frequencies between the Malaysian CD patients and controls. This shows that the DLG5 (4136 C/A) polymorphisms have a high degree of association with the onset of CD. The C allele presented almost thrice more often than the A allele in patients, which implies that the C allele may be one of the predisposing factors for the onset of CD. In contrast, C and A alleles are distributed equally in the control individuals. In the present study, the homozygous C genotype was observed to be more common in both patients and control individuals than homozygous A and the heterozygous C/A genotypes. Conversely, the homozygous A genotype was observed to be significant in the controls while the homozygous A and heterozygous C/A genotypes exist evenly in CD patients. These three genotypes are present in similar ratios in the controls.

**DLG5_e26 polymorphisms**

Figure 2 shows the banding patterns observed in the study of DLG5_e26 polymorphisms while the genotype and allelic frequencies for DLG5_e26 polymorphisms are shown in Table 3. The presence of the insA allele was significant in CD patients, which was thrice more frequent than the delA allele. The homozygous insA genotype was significant in CD patients while in contrast, the heterozygous insA/delA genotype was significantly associated with the normal control group. Hence, it is not surprising that the homozygous insA genotype presented more frequently in CD patients while heterozygous insA/delA genotype was higher in controls, as compared to the other two genotypes.

**SLC22A5 (-207 G/C) polymorphisms**

Figure 3 shows the separation of PCR-RFLP products for the investigation of SLC22A5 (-207 G/C) polymorphisms. There was no significant association between allelic frequencies and the onset of CD (Table 4). The distribution of genotype frequencies between CD patients and control individuals also did not reveal a significant correlation. The homozygous G genotype was observed most frequently in CD patients, and was found in 42.50% of all CD patients in this study. In contrast, heterozygous G/C was more common in the control individuals (38.00%).

**DISCUSSION**

**DLG5 (4136 C/A) polymorphisms**

The DLG5 (4136 C/A) polymorphisms in exon 23 of the DLG5 gene result in the amino acid substitution

---

**Table 2. The distribution of genotype and allele of DLG5 (4136 C/A) polymorphisms in Malaysian Crohn’s disease (CD) patients and controls**

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Frequency (n [%])</th>
<th>CD patients (N = 80)</th>
<th>Control (N = 100)</th>
<th>χ² (P value)</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genotype</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C/C</td>
<td>47 (58.75)</td>
<td>36 (36.00)</td>
<td>9.2571 (0.0023)</td>
<td>2.5320 (1.3840–4.6324)</td>
<td></td>
</tr>
<tr>
<td>C/A</td>
<td>18 (22.50)</td>
<td>30 (30.00)</td>
<td>1.2784 (0.2582)</td>
<td>0.6774 (0.3442–1.3331)</td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>15 (18.75)</td>
<td>34 (34.00)</td>
<td>5.2172 (0.0224)</td>
<td>0.4480 (0.2230–0.8998)</td>
<td></td>
</tr>
<tr>
<td>Allele</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>112 (70.00)</td>
<td>102 (51.00)</td>
<td>13.3105 (0.00026)</td>
<td>2.1961 (1.4209–3.3942)</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>48 (30.00)</td>
<td>98 (49.00)</td>
<td>0.4461 (0.2881–0.6908)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OR, odds ratio; CI, confidence interval.
P1371Q, that is, the substitution of proline by glutamine. The P1371 is located before the fourth PDZ domain in the MAGUK protein and is the third proline which contains a PXXP sequence motif for binding to an SH3 domain. PDZ domains are crucial for the localization of MAGUK and other proteins that interact with them to bind to the membrane domains of epithelial cells. PDZ domains also play a role in helping MAGUK proteins to bind to C-terminal recognition sequences.\textsuperscript{16} We postulate that this amino acid change will hinder the normal function of the PDZ domain as it precedes the fourth PDZ domain. In addition, the PDZ domain cannot bind to the SH3 domain without the PXXP sequence motif, which may result in a defective MAGUK protein. All these domains are involved in supporting the \textit{DLG5} gene to maintain the epithelial structure.\textsuperscript{31} Therefore, we suggest that this polymorphism may play a role in deterring the normal functions of the epithelial barrier in the gastrointestinal tract, which can then lead to the inflammation of the affected area.

In this study, we found significant associations in the \textit{DLG5} (4136 C/A) polymorphism in the cohort of Malaysian CD patients. The C allele was observed to be the predominant allele in CD patients. We hypothesize that the \textit{DLG5} transcription process is altered in the presence of C allele(s). The gene expression

\begin{table}
\centering
\caption{The distribution of genotype and allele of \textit{DLG5} \textsubscript{e26} polymorphisms in Malaysian Crohn’s disease (CD) patients and controls}
\begin{tabular}{lccc}
\hline
\textbf{DLG5\textsubscript{e26}} & \textbf{CD patients (N = 80)} & \textbf{Control (N = 100)} & \textbf{$\chi^2$ (P value)} & \textbf{OR (95% CI)} \\
\hline
\textbf{Genotype} & & & & \\
\text{insA/insA} & 47 (58.75) & 33 (33.00) & 11.9351 (0.0006) & 2.8916 (1.5714–5.3209) \\
\text{insA/delA} & 26 (32.50) & 58 (58.00) & 11.6116 (0.0007) & 0.3487 (0.1888–0.6441) \\
\text{delA/delA} & 7 (8.75) & 9 (9.00) & 0.0034 (0.9533) & 0.9696 (0.3446–2.7286) \\
\textbf{Allele} & & & & \\
\text{insA} & 120 (75.00) & 124 (62.00) & 6.8785 (0.0087) & 1.8387 (1.1633–2.9062) \\
\text{delA} & 40 (25.00) & 76 (38.00) & 0.5439 (0.3441–0.8597) & \\
\hline
\end{tabular}
\end{table}

\begin{table}
\centering
\caption{The distribution of genotype and allele of \textit{SLC22A5} (-207 G/C) polymorphisms in Malaysian Crohn’s disease (CD) patients and controls}
\begin{tabular}{lccc}
\hline
\textbf{SLC22A5} (-207 G/C) & \textbf{CD patients (N = 80)} & \textbf{Control (N = 100)} & \textbf{$\chi^2$ (P value)} & \textbf{OR (95% CI)} \\
\hline
\textbf{Genotype} & & & & \\
\text{G/G} & 34 (42.50) & 33 (33.00) & 1.7165 (0.1901) & 1.5007 (0.8166–2.7577) \\
\text{G/C} & 29 (36.25) & 38 (38.00) & 0.0582 (0.8093) & 0.9278 (0.5046–1.7058) \\
\text{C/C} & 17 (21.25) & 29 (29.00) & 1.4031 (0.2362) & 0.6606 (0.3320–1.3146) \\
\textbf{Allele} & & & & \\
\text{G} & 97 (60.63) & 104 (52.00) & 2.6815 (0.1015) & 1.4212 (0.9325–2.1659) \\
\text{C} & 63 (39.37) & 96 (48.00) & 0.7036 (0.4617–1.0723) & \\
\hline
\end{tabular}
\end{table}

Figure 3. Polymerase chain reaction (PCR)–restriction fragment length polymorphism (RFLP) analysis of \textit{SLC22A5} (-207 G/C) polymorphisms on agarose gel. Lane 1, 50 bp DNA marker; Lane 2, undigested PCR product; Lane 3, heterozygous G/C; Lane 4, homozygous G/G; Lane 5, homozygous C/C; Lane 6, DNA blank.
products from the C allele may influence an individual’s susceptibility to CD. In contrast, the homozygous A genotype is observed to be significant in healthy controls. Hence, it is possible that the A allele may have a protective role against the onset of CD. Furthermore, we also observed higher frequencies of heterozygous C/A genotypes in the controls. Perhaps the A allele is the dominant allele that masks the effects of its C counterpart.

Similar observations were obtained in a recent study carried out by Weersma et al. in which DLG5 (4136 C/A) was found to have a significant association with the development of CD in the Dutch population. The C allele was found to exist more frequently in Dutch CD patients than in the controls. The study also reported that the C allele was responsible for the structuring behavior of CD. However, the authors suggested that DLG5 (4136 C/A) is a low-penetrant variant. Therefore, they carried out an analysis of gene to gene interaction effects with the other risk alleles, including the C allele in IBD5 (rs2522027), the G allele in ATG16L1 (rs2241880) and the G alleles in IL23R (rs11209026). As there was a significant association of these gene to gene interactions with the onset of CD and increased OR for CD, it was therefore suggested that these gene to gene interaction may increase the risk and severity of CD. It is reported that individuals carrying an increased number of risk alleles will have a higher risk of developing CD.

Yamazaki et al. reported that DLG5 (4136 C/A) polymorphisms were also causative in an analysis of CD patients from Japan in 2004. This is due to the higher frequency of participants with a homozygous C genotype, while the homozygous A genotype was rare. In contrast, there are studies by other researchers showing no relation between DLG5 (4136 C/A) polymorphisms and CD. In 2005, Török et al. did not observe any significant results from his study in Germany. They reported higher homozygous C genotypes among the German population. Another study carried out in Germany by Stoll et al. also reported no significant association between the CD patients and normal controls (P = 0.2743). Thus far, the DLG5 (4136 C/A) polymorphism has been shown to be significantly associated with CD in Dutch and Canadian groups, but not in New Zealand Caucasians and Germans. These findings agree with the postulation that ethnic variety may contribute to the allelic and genotypic frequencies in different populations, especially between Western (Caucasian) and Asian groups.

### DLG5_e26 polymorphisms

Haplotypes A tagged DLG5_e26 is known to be under-transmitted in individuals with CD. It is a genetic variant of the DLG5 gene that has shown protective function against the onset of CD. In our study, the homozygous insA genotype was observed to be significant in CD patients. In contrast, the heterozygous insA/delA genotype was observed to be significant in the controls. Therefore, we propose that the defense function may due to a complex interaction of both insA and delA alleles in the heterozygous genotype. Similar results were reported in a study carried out in 2004 by Stoll et al. on IBD patients in Germany. They reported that the insertion of adenine was significantly associated with the development of CD (P = 0.0075).

Weersma et al. conducted a study on the allelic association of DLG5_e26 polymorphisms with CD and observed no significant association. Prior to that, Török et al. had shown that polymorphisms in the DLG5_e26 did not contribute to the risk of CD development nor did it protect individuals from getting the disease. They also observed that the heterozygous insA/delA genotype was the major genotype in the German population. On the other hand, the homozygous insA genotype is the major genotype in the Malaysian population. The differences in results obtained could therefore be due to ethnical divergence. There is limited information about the relationship of the DLG5_e26 polymorphism with the onset of CD in Asian groups. Yamazaki et al. conducted a study to investigate the distribution of DLG5_e26 polymorphisms in the Japanese population, but observed no significant relationship between CD patients and normal controls either. However, they reported that homozygous insA genotype is the major genotype in the Japanese population. It is obvious that allelic and genotypic frequencies of the DLG5_e26 polymorphisms vary in different parts of the world. Ethnic divergence may affect the susceptibility of an individual to develop CD, but as different populations may share similar phenotypes, they may not necessarily share identical predisposing variants.

### SLC22A5 (-207 G/C) polymorphisms

In the present study, no significant association was observed between the allelic and genotypic frequencies of these polymorphisms and susceptibility to CD. The G allele presented more frequently than the C allele in both CD patients and controls. A similar
analysis was carried out by Martínez et al. on a Spanish group. They reported an insignificant association between this polymorphism and CD as the G allele presented in more than half of the study group while the average frequency of the C allele was 46.8%. The homozygous G and homozygous C genotypes were observed evenly in both the CD patients and controls, while the heterozygous G/C genotype was more common in both groups, i.e. 50.85% and 46.30%, respectively.

Interestingly, Li et al. reported that these genetic polymorphisms were completely absent in the Chinese population. The C allele was not presented in either CD patients or controls. Thus, they concluded that this genetic variant did not play a part in the development of CD. In another study carried out in Japan, no evidence was found to support the association of SLC22A5 (-207 G/C) polymorphisms and CD. These polymorphisms were found to have a significant association in some groups, such as Greece, Germany and Italy, although it was reported to be insignificant in other groups such as in Spain, China, Japan and Malaysia. Our study was similar to those reported in other Asian populations.

In conclusion, we reported a significant association between polymorphisms in DLG5 polymorphisms with CD in the Malaysian population. Although it is generally accepted that ethnic divergence not only affects the susceptibility of an individual to developing CD; there is now also increasing evidence that different predisposing variants are seen in Asian and western populations. However, the findings of this study and our previous study suggest that there is still an overlap in the putative genes involved in the pathogenesis of CD between these broad groups. And as more data emerges, we will move ever closer to fully understanding this challenging condition.

ACKNOWLEDGEMENT
The study was supported by University of Malaya Research Grant (RG274/10HTM).

REFERENCES


