Disease Markers 27 (2009) 193–201
DOI 10.3233/DMA-2009-0662
IOS Press

Genetic association of DLG5 R30Q with familial and sporadic inflammatory bowel disease in men

Z. Lin[a,*], L. Poritz[a], A. Franke[b], T.Y. Li[c], A. Ruether[b], K.A. Byrnes[a], Y. Wang[d], A.W. Gebhard[a], C. MacNeill[e], N.J. Thomas[f], R. Wu[g], S. Schreiber[b,h] and W.A. Koltun[a,*]
[a]*Department of Surgery, The Pennsylvania State University College of Medicine, Hershey, Pennsylvania, USA*
[b]*Institute for Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany*
[c]*Total Health Management and Technology, Inc., Philadelphia, Pennsylvania, USA*
[d]*Department of Molecular and Cellular Physiology, The Pennsylvania State University College of Medicine, Hershey, Pennsylvania, USA*
[e]*Department of Obstetrics and Gynecology, The Pennsylvania State University College of Medicine, Hershey, Pennsylvania, USA*
[f]*Department of Pediatrics, The Pennsylvania State University College of Medicine, Hershey, Pennsylvania, USA*
[g]*Center for Statistical Genetics, The Pennsylvania State University College of Medicine, Hershey, Pennsylvania, USA*
[h]*Department of General Internal Medicine, Christian-Albrechts-University, Kiel, Germany*

Abstract. *Background:* The association of DLG5 R30Q with IBD has been replicated in several populations, but is not statistically significant in others. We studied the incidence of DLG5 alleles in a population of IBD patients from Pennsylvania.
Methods: DLG5 R30Q (rs1248696) and G1066G (rs1248634) were analyzed with PCR-based RFLP methods in a total of 521 subjects, that included 105 individuals with IBD and 139 without IBD from a familial IBD registry, 107 with sporadic IBD, and 170 unrelated healthy controls. R30Q was further analyzed with SNPlex$^{\mathrm{TM}}$ Genotyping System in 473 samples.
Results: RFLP genotyping data showed that, DLG5 R30Q was significantly associated with IBD overall ($p = 0.006$), and separately with CD ($p = 0.009$) and UC ($p = 0.024$). The association of R30Q with IBD was entirely due to a male-associated effect (male vs female $p = 0.015$ vs 0.241 (IBD), $p = 0.024$ vs 0.190 (CD), and $p = 0.019$ vs 0.575 (UC)). The frequency of the A allele carriage was elevated in both affected and unaffected members in the familial IBD cohort compared to healthy controls ($p = 0.037$). In the family pedigrees, we observed differences in the expression of IBD in individuals carrying the A allele between families.
Conclusions: In the studied population, DLG5 R30Q was associated with all forms of IBD. An elevated presence of the R30Q variant was observed in all members of a familial IBD registry. This association of the R30Q variant with IBD was male-specific.

Keywords: DLG5 R30Q, familial IBD registry, genetic association, inflammatory bowel disease, male-specific

1. Introduction

Inflammatory bowel disease (IBD), consisting of two major subgroups Crohn's disease (CD) and ulcerative colitis (UC), is a heterogeneous disease of intestinal inflammation, whose cause and clinical expression are affected by both genetic and environmental factors [1–3]. Evidence from both animal models and human studies,

*Corresponding authors: Walter A. Koltun MD, Department of Surgery, The Pennsylvania State University College of Medicine, 500 University Drive, H 137, Hershey, Pennsylvania 17033. Tel.: +717 531 5164; Fax: +717 531 0646; E-mail: wkoltun@psu.edu; Zhenwu Lin PhD, Department of Surgery, The Pennsylvania State University College of Medicine, 500 University Drive, H 137, Hershey, Pennsylvania 17033. Tel.: +717 531 6196; Fax: +717 531 0646; E-mail: zlin@psu.edu.

ISSN 0278-0240/09/$17.00

including twin [4,5] and family studies [6,7]. indicate that the development of IBD has a strong genetic component [8]. The first IBD-associated gene conferring susceptibility to CD, NOD2/CARD15, was identified on chromosome 16q12 in 2000 [9–11], Additional associations of DLG5 [12] and gene clusters containing OCTN1 and OCTN2 [13] have been described. Recently, using genome-wide association technique, additional IBD associated genes/regions have been identified. These include *TNFSF15* [14], *IL23R* [15], *ATG16L1* [16,17], the $5p13.1$ region [18], and IRGM, NKX2-3 and PTPN2 [19,20].

DLG5 is a member of the membrane associated guanylate kinase family that plays a role in the formation of cell junctions, maintenance of cell shape and clustering of channel proteins at the cell surface [21–23]. The DLG5 variant 113G→A (R30Q) leads to the change of amino acid 30 in exon 3 from Arginine (G allele) to Glutamine (A allele) in the DLG5 protein, and may affect protein function. Stoll et al., [12] identified an IBD risk-associated DLG5 haplotype D that is uniquely distinguished by the 113A variant of R30Q. The association of R30Q with IBD has been replicated in other studies from several different patient populations [24–26]. Friedrichs et al., found a male-specific association of the R30Q variant with CD, with no association of the R30Q variant with women, nor with the entire sample of men and women [27]. However, the association of the R30Q variant with IBD failed to be replicated by several other groups with samples from several different countries, including Germany, Scotland, Belgium, and Hungary [28–36]. A US based population of IBD patients has been also studied [17, 37]. Biank et al found a significant negative association between R30Q in female children [37].

The purpose of the present study was to evaluate the possible association of the R30Q allele of the DLG5 gene in a population of patients from Central Pennsylvania in the US using a familial registry. There has been no study to date, to our knowledge, investigating DLG5 R30Q in a patient sample from a familial IBD registry.

The Central Pennsylvania area is a relatively unique rural/suburban area. Historically it was largely populated by immigrants from Western Europe (Germany, the Netherlands) and the United Kingdom, and thus presents itself as having a patient population with a relatively homogeneous genetic background. A familial registry drawing on this population was established in 1999 at the Milton S. Hershey Medical Center recruiting families having at least two family members with IBD. Inclusion in the registry involved the harvesting and EBV immortalization of lymphocytes with clinical characterization of disease from both IBD affected and unaffected individuals within the family. The samples studied in the present project also included patients with sporadic IBD. In addition to the nonsynonymous DLG5 SNP R30Q, we included another DLG5 synonymous SNP G1066G in this study. The G1066G SNP does not alter the amino acid residue of the protein, and likely has no impact on DLG5 expression and function.

2. Materials and methods

2.1. Study participants and human tissues

A total of 716 samples were studied including:

1) Familial IBD patients from the Milton S Hershey Familial IBD Registry. A total of 244 individuals from 58 families were studied, 106 individuals were affected with IBD (CD: 59 and UC: 47) and 139 were without IBD. These included 124 males and 120 females ranging in age from 16–92 (mean 51) years old. Blood was collected from the participants and used to derive B lymphocyte cell lines by EB virus transformation [38].

2) Sporadic IBD patients ($n = 107$) from the Milton S. Hershey Medical Center consisted of 57 patient with CD and 50 with UC. The age range was 22–66 (mean 48) years. There were 59 males, and 48 females. DNA in this group was obtained from intestinal tissue harvested at the time of surgery.

3) Unrelated healthy controls ($n = 365$). The age range was 15–81 years and the average age was 37. Most of the DNA was isolated from blood, but 60 was from lung tissues that were obtained from the Gift of Life Donor Program (Philadelphia, PA).

Genotype analysis using the PCR-based RFLP-cRFLP method was performed on 521 samples including all 212 IBD patients, 170 unrelated controls, and 139 non-IBD individuals from the familial IBD registry. For genotype analysis with the SNPlex$^{\mathrm{TM}}$ Genotyping System 473 samples were analyzed, including most individuals of the familial IBD registry, sporadic IBD, and additional newly recruited unrelated healthy controls.

All the human tissues described above were approved by the Human Subjects Protection Offices of The Pennsylvania State University College of Medicine, and were undertaken with the understanding and written consent of each subject.

Table 1
Primers and restriction enzymes used in this study

SNP	Primer	Sequences (5' to 3')	RE	PCR primers	Restriction fragment (bp)
G113A	3DLG5r	TCATCAGCCGCAGCTGAATG	Msp 1	20DLG5f and 3DLG5r	A allele: 103bp; G allele: 51 + 52bp
(R30Q)	20DLG5f	TTCCACAGGCACTACTGG GA			
	29DLG5f	ACCACCCCTCCTCACTGCCC	BseY 1	29DLG5f and 30DLG5r	A allele: 120 + 20bp; G allele: 140bp
	30DLG5r	ACCTCTTGTCAAAGGCCGTG			
C3222T	8DLG5r	TGGAGCTCACAGTGGTGCTG	Rsa 1	11DLG5f and 8DLG5r	C allele: 66bp; T allele: 46 + 20bp
(G1066G)	11DLG5f	CTCACCAAGCCCATTTCTATG			
	12DLG5r	GGTCAAACTCCGGGGAACATT	HpyCH4 V	11DLG5f and 12DLG5r	C allele: 24 + 20bp; T allele: 44bp

Underline letter: mismatched nucleotide. RE, restriction enzyme. The last letter in the name of primers; f, forward; r, reverse.

2.2. DNA isolation

Genomic DNA used in this study was isolated from the tissues and cells. For the B cell lines, cells were cultured in RPMI containing 12% FBS using standard sterile culture technique. Genomic DNA was isolated with the Gentra Systems kit (Minneapolis, MN). The final DNA pellet was suspended in 10 mM Tris buffer (pH 8.0). For intestinal tissue, blood, and lung tissue, genomic DNA was isolated with QIAamp DNA Mini Kit (Qiagen Inc. Valencia, CA) from pulverized tissue powder according to manufacturer's instruction. The DNA was eluted from the QIAamp column with distilled water (dH_2O). DNA concentration was measured with a Nanodrop ND-1000 spectrophotometer (NanoDrop Technology, Wilmington, DE), and stored at $-80°$C until use.

2.3. Genotyping with pcr-based rflp-crflp method

The two DLG5 variants R30Q (113G→A) (rs12486-96) and G1066G (3222C→T) (rs1248634) were genotyped using PCR-based restriction fragment length polymorphism (RFLP) and converted RFLP (cRFLP) methods [39].

The DLG5 SNP R30Q (113G→A) is located in exon 3 (CCC A/G GC).

The G allele, CCGG, is an Msp I recognition site, but the A allele, CCAG, is not. A 103bp DNA fragment was amplified from genomic DNA with PCR primers 3DLG5r and 20DLG5f. All the primers used in this study are listed in Table 1. After the PCR products were digested with Msp I, the PCR products of the A allele remained 103bp in length, while that of the G allele became 51 and 52bp long (Table 1 and Fig. 1A).

To confirm the results obtained from Msp I RFLP, the BseY I RFLP method was used. The primers 29DLG5f and 30DLG5r were used for PCR amplification. In BseY I RFLP, the PCR products of the A allele (CCCAGC) were cut by the BseY I enzyme (yielding fragments of 120 and 20 bp), but that of the G allele (CCCGGC) were not (Table 1 and Fig 1B).

The DLG5 SNP G1066G (3222C→T) is located in exon 16 (GG C/T AC). The 3222T allele, GTAC, is a Rsa I recognition site, but the 3222C allele, GCAC, is not. A 66 bp DNA fragment was amplified with PCR primers 11DLG5f and 8DLG5r. After PCR product digestion with Rsa I, PCR products of the C allele remained 66 bp, while that of the T allele became 46 and 20 bp (Table 1 and Fig. 1C).

The genotyping results from Rsa I RFLP were confirmed by the HpyCH4 V cRFLP method. To convert the G1066G (3222C→T) to a HpyCH4 V recognition site, a mismatched nucleotide T to G was introduced at the 2nd nucleotide of 3' end of primer 11DLG5f. By incorporating the mismatched T into the PCR products, the SNP C3222T was converted to TG C/T A (the original sequence is GG C/T A) where the C allele, TGCA, was cut by HpyCH4Y I (fragments 24 and 20 bp), and the T allele, TGTA was not (fragment 44 bp) (Table 1 and Fig. 1D).

For each RFLP/cRFLP analysis, more than 10% of the samples were randomly selected and repeated.

One hundred ng of DNA was used for the PCR in a 30 μl volume. The PCR profile was as follows: 95°C for 2 min, 5 cycles of 95°C for 30 sec, 50°C for 1 min, and 72°C for 1 min, then 30 cycles of 95°C for 30 sec, 58°C for 1 min, and 72°C for 1 min, followed by a final extension step at 72°C for 4 min.

PCR products (5 μl) were digested with an appropriate restriction enzyme according to manufacturer's instructions.

The digested PCR products were separated on PAGE (8% for R30Q, 10% for G1066G). The gel photograph is partly shown in the Fig. 1.

2.4. Genotyping with SNPlex$^{\mathrm{TM}}$ Genotyping System

DLG5 R30Q was included in another genetic association study using the SNPlex$^{\mathrm{TM}}$ Genotyping System (Applied Biosystems, Foster City, CA, USA) as

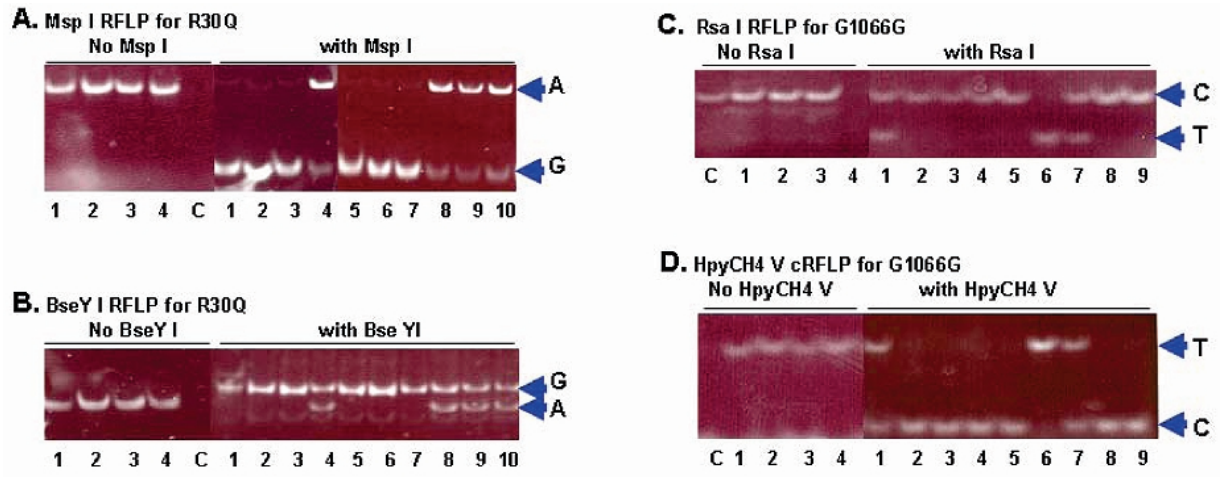

Fig. 1. Genotype analysis of the DLG variants R30Q and G1066G. PCR amplification and restriction enzyme digestion were described in the Material and Methods. After the restriction enzyme digestion, PCR products were loaded on 8% PAGE for R30Q and 10% PAGE for the G1066G to separate allelic products, then stained with ethidium bromide and photographed under UV light. The patterns of allelic products of R30Q on the 8% PAGE with Msp I RFLP method were shown in Panel **A,** and with BseY I RFLP method were shown in Panel **B**. The patterns of allelic products of G1066G on the 10% PAGE with Rsa I RFLP method are shown in Panel **C**, and with HpyCH4 V cRFLP method is shown in Panel **D**. In panels B, C, and D, the 20bp fragments run out of the gel and thus were not seen. Numbers 1 to 10 denote the respective DNA sample; C was the PCR control without DNA. The DNA fragments for A and G alleles of R30Q and for C and T alleles of G1066G were indicated by arrows.

described previously [16,40,41]. In brief, 100 ng of genomic DNA were used for genotyping using the SNPlex^{TM} chemistry (Applied Biosystems) on an automated platform with TECAN Freedom EVO and 384well TEMO liquid handling robots. Genotype assignments were confirmed by visual inspection with the Genemapper® software v4.0 (Applied Biosystems). The samples used for SNPlex^{TM} Genotyping were the same as for RFLP-cRFLP methods, but without sporadic CD samples, but including more unrelated healthy controls.

The genotype concordance for both SNPs under study was 100% between all the samples ($n = 521$) for the two different RFLP methods, between two repeats in each RFLP analysis (for each RFLP of each SNP, sample $n = 95$), and between the two genotyping methods of PCR-based RFLP and the SNPlex^{TM} genotyping system (samples $n = 339$ for the technical validation).

2.5. Statistical analysis

The genotype data of the 521 samples using the PCR-based RFLP method was used for statistic analysis using SAS software version 9.1.3 (SAS Institute; Cary, NC). Pearson's X^2-test with one degree of freedom for allelic association were performed using Haploview. We also calculated Genotype-based Odds Ratio (OR) using Fisher's contingency tables and tested association similarly. The difference was considered as significant when $p < 0.05$.

3. Results

3.1. Genetic association of the DLG5 R30Q with IBD, CD, and UC

The DLG5 R30Q was almost entirely in the form of the heterozygous AG (only one was AA in the total of 521 studied samples. The frequency of A allele carriage in healthy controls was only 13%, much less common than the G allele (87%). However, the carriage of the A allele in the sporadic IBD patients was 20.2%. In both familial IBD ($n = 105$) and sporadic IBD ($n = 107$), the frequency of A allele carriage was 19.1%. The statistical analysis showed that the R30Q variant was significantly associated with IBD (OR = 2.131, 95% CI 1.233–3.684, $\chi^2 = 7.541$, $p = 0.006$) (Table 2).

Further analysis of the R30Q allele separating the CD and UC patients indicated a significant association of R30Q with each of CD and UC in the studied populations (familial and sporadic) (Table 2): for CD patients, OR = 2.242, 95% CI 1.213–4.144, $\chi^2 = 6.843$, and $p = 0.009$; and for UC patients, OR = 2.091, 95% CI 1.094–3.996, $\chi^2 = 5.112$, and $p = 0.024$. Table 2 also shows that there was no association of the G1066G with CD (OR = 1.192, 95% CI 0.743–1.913,

Table 2
Association of DLG5 variants R30Q and G1066G with IBD in sporadic IBD patients and IBD patients from the familial IBD registry

SNP	Genotype	Control (n)	Disease n (Sn, Fn)§	OR	95% CI	X^2	p value
A. For IBD patients*							
For IBD							
R30Q	GA	22	51 (29, 22)	2.131	1.233–3.684	7.541	0.006
	GG	148	161 (78, 83)				
G1066G	CT + TT	79	112 (65, 47)	1.290	0.861–1.933	1.526	0.217
	CC	91	110 (42, 58)				
For CD							
R30Q	GA	22	29 (16, 13)	2.242	1.213–4.144	6.843	0.009
	GG	148	87 (42, 45)				
G1066G	CT + TT	79	59 (33, 26)	1.192	0.743–1.913	0.533	0.466
	CC	91	57 (25, 32)				
For UC							
R30Q	GA	22	23 (14, 9)	2.091	1.094–3.996	5.112	0.024
	GG	148	74 (36, 38)				
G1066G	CT + TT	79	54 (33, 21)	1.447	0.876–2.388	2.091	0.148
	CC	91	43 (17, 26)				
B. For family members¶							
		(n)	n (Wn, On)¥				
R30Q	GA	22	51 (22, 29)	1.778	1.032–3.062	4.372	0.037
	GG	148	193 (83, 110)				
G1066G	CT + TT	79	113 (61, 53)	0.994	0.671–1.471	0.001	0.975
	CC	91	131 (44, 87)				

* Patients include both sporadic IBD patients and IBD patients from the familial IBD registry.
§ Sn is the number of sporadic IBD, while Fn is the number of familial IBD from familial IBD registry.
¶ Members include both with and without IBD that are collected in the familial IBD registry.
¥ Wn is the number of familial members with IBD while On is the number of familial members without IBD from the familial IBD registry.

$\chi^2 = 0.533$, and $p = 0.466$), UC (OR = 1.447, 95% CI 0.876–2.388, $\chi^2 = 2.091$, and $p = 0.148$), and thus not with IBD patients overall.

It is very interesting to note that in all family members (both IBD affected ($n = 105$) and unaffected ($n = 139$) with IBD) the A allele carriage was as high as 20.9%. Compared to unrelated healthy controls (13%), the frequency of DLG5 A allele carriage in these family members were very high, similar to the sporadic IBD patients (20.9%). This was statistically significant (OR = 1.778, 95% CI 1.032–3.062, $\chi^2 = 4.372$, and $p = 0.037$). In these family members, more than half were not affected with IBD (105 individuals were affected with IBD and 139 individuals without IBD), suggesting that DLG5 R30Q is a risk factor for IBD development, but not sufficient for disease expression. For the SNP G1066G in these families, as expected, there was no association with IBD (OR = 0.994, 95% CI 0.671–1.471, $\chi^2 = 0.001$, and $p = 0.975$) (Table 2).

G1066G is another DLG5 variant. This substitution of T to C occurs at the third nucleotide of the amino acid codon 1066 glycine, resulting in no amino acid change (synonymous). Thus the G1066G variant is presumed to have no direct impact on DLG5 protein function. The genotyping results (Table 2) indicated that the frequency of carriership of T (i.e. CT and TT) was 53% among all IBD patients and 47% in healthy controls. Statistical analysis showed that there was no association of the G1066G with IBD (OR = 1.290, 95% CI 0.861–1.933, $\chi^2 = 1.526$, $p = 0.217$).

3.2. *Male-specific association of the DLG5 R30Q with IBD, CD and UC*

Some epidemiological data suggests gender related differences in CD incidence. We further examined whether gender, as a modifier, affected the association of the R30Q with IBD. The results indicated that the association of the R30Q was entirely due to a male-specific association with IBD ($p = 0.015$ in men vs $p = 0.241$ in women), with CD ($p = 0.024$ in men vs $p = 0.190$ in women), and with UC ($p = 0.019$ in men vs $p = 0.575$ in women) (Table 3). For the SNP G1066G, there was no male-specific association

Table 3
Distribution of the 113A variant of DLG5 gene in men and women

	All				Males				Females			
SNP	OR	95% CI	X^2	p value	OR	95% CI	X^2	p value	OR	95% CI	X^2	p value
For IBD												
R30Q-A	2.131	1.233–3.684	7.541	0.006	3.333	1.208–9.193	5.867	0.015	1.525	0.751–3.098	1.375	0.241
G1066G-T	1.290	0.865–1.933	1.526	0.217	1.375	0.694–2.725	0.836	0.361	1.383	0.814-2.550	1.444	0.230
For CD												
R30Q-A	2.242	1.213–4.144	6.843	0.009	3.323	1.126–9.805	5.079	0.024	1.724	0.759–3.913	1.720	0.190
G1066G-T	1.192	0.743–1.913	0.533	0.466	1.269	0.592–2.720	0.375	0.541	1.281	0.677–2.426	0.580	0.446
For UC												
R30Q-A	2.091	1.094–3.996	5.112	0.024	3.583	1.187–10.82	5.534	0.019	1.300	0.519–3.258	0.315	0.575
G1066G-T	1.447	0.872–2.388	2.091	0.148	1.456	0.658–3.222	0.862	0.353	1.661	0.833–3.313	2.093	0.148

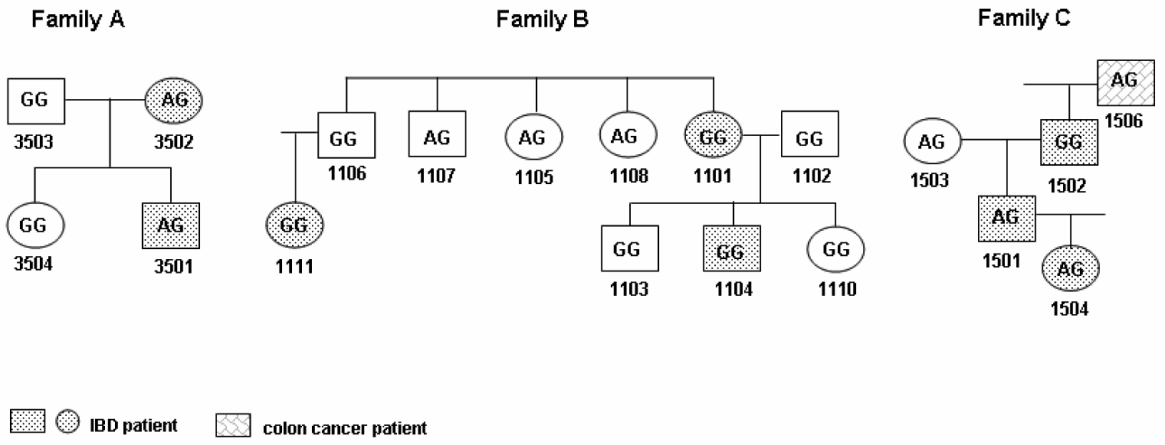

Fig. 2. Clinical expression of IBD in families carrying the A allele of DLG5 of R30Q Individuals affected with IBD are shown with filled dots, affected with colon cancer with filled lines, and non-diseased without filling. Number was sample number. AG and GG were genotypes for the R30Q variant: A = A allele; G = G allele.

with IBD ($p = 0.361$ in men vs 0.230 in women), CD ($p = 0.541$ in men and 0.446 in women), nor UC ($p = 0.353$ in men and 0.148 in women) (Table 3). This indicates that the R30Q variant conferred a much higher risk of IBD in men than women. The results presented here support the previous finding that R30Q is a male-specific association in CD [27], and further reveals that the male-specific association of the R30Q variant is also seen in UC, and thus in IBD overall in the studied populations.

3.3. Clinical expression of IBD between families carrying the A allele of DLG5 R30Q

A varied clinical expression or phenotype of the R30Q allele was observed among different families carrying the A allele. Figure 2 shows 3 typical examples. In family A, both family members carrying the A allele were affected with IBD, while the other 2 members without the allele were not. Conversely, in family B, none of the three family members carrying the A allele was affected with IBD, while none of the three family members affected with IBD carried the A allele. In family C, the clinical expression of IBD varied in the 4 family members carrying the allele. Two of the carriers were affected with IBD, one was affected with colon cancer, and the other had no disease symptoms (Fig. 2). These observations suggest that although the R30Q variant was significantly associated with IBD on a statistical basis, clearly it was only a weak factor and was neither sufficient nor necessary to cause IBD. The clinical expression of the A allele, in so far as IBD is concerned, is likely affected by other genetic and environmental factor(s).

4. Discussion

The present report shows that the R30Q variant of the DLG5 gene was associated with both CD and UC in the studied Pennsylvania population. Further analysis indicated that the association was entirely due to a

male-specific association with IBD, CD, and UC. We also observed an elevated presence of the R30Q variant in both those family members affected and those not affected with IBD in the studied familial IBD registry, suggesting that the R30Q variant is a risk factor for developing IBD. The evidence, that the DLG5 nonsynonymous SNP R30Q showed an association with IBD while the DLG5 synonymous SNP G1066G did not, is supportive that DLG5 R30Q may be a functional SNP in IBD pathogenesis, although this needs to be functionally studied.

Since Stoll et al., reported that the R30Q was associated with CD patients, several groups have studied its effect on CD and IBD in different populations. The association of the R30Q variant with IBD has been supported in some studies [12,24,25,27], but failed to be replicated in others [28–34].

The varied association of the DLG5 gene with IBD may reflect the heterogeneity of the disease process and implies that DLG5 is a moderate/weak factor for IBD. IBD is a multigenic complex disease and is affected by environmental factors. In a hypothetical model, genetic mutations could be classified as having either a major or minor effect on disease expression, while other genes could be modifiers of disease [42]. DLG5 may be a minor factor, whose disease-causative effects could be easily modified in either a positive or negative way by other genes or other influences including gene-environment interactions (e.g. epigenetic modification of genetic elements). The clinical phenotype expression differed between families carrying the A allele and the gender-specific association of R30Q with IBD support this assertion.

It will be difficult to definitively elucidate the role of DLG5 in the pathogenesis of IBD, because DLG5 is only one of many potential susceptibility genes, each of which could play a larger or lesser role in disease pathogenesis, depending on the relative presence or absence of other genetic and environmental determinants. Based on the present study using familial IBD pedigrees, further study on the selected families and individuals where significant associations do exist will facilitate mechanistic study of DLG5 in IBD pathogenesis.

To evaluate the role of the DLG5 gene in IBD pathogenesis in this study, we used well-characterized patients with a relatively homogeneous ethnic background. Our familial registry consists of 99% Caucasian and 73% of individuals are descendants or immigrants from the UK, Germany or the Netherlands. Using the same population, we have previously observed a high frequency of NOD2 mutations which were found in approximately 50% of studied families [43]. Compared to previous studies looking at DLG5 where case-control studies were used or parent-offspring trio samples, the present study is the first one using a familial IBD registry in an US based environment. The lack of relative ethnic diversity in our familial patient registry may allow the discovery of or association of susceptibility genes that are more difficult to define in more heterogeneous populations. The result reported here may be unique to this relatively less heterogeneous group of Pennsylvania patients. The studied familial IBD registry may provide family pedigrees to allow future study of the mechanism of disease promotion by the DLG5 R30Q.

The present study reveals a male-specific association of the R30Q variant with IBD overall, as well as with CD and UC separately. Male-specific linkage has been found in the major histocompatibility region of chromosome 6p (IBD3) in affected families of both CD and UC. Regions on chromosome 11, 14, and 18 also have shown strong evidence of linkage in male-affected families but not in female-affected families [44]. A male-specific association of the R30Q variant with CD has been recently reported by Friedriches et al., [27]. These observations raise very interesting questions including what is the "gender factor" and how does it affect IBD development? Whether the gender factor inhibits a protective effect associated with the X chromosome, or conversely, facilitates a disease promoting gene that is associated with the Y chromosome, or whether endocrine tissues play a role, are currently unknown. Recently, Biank et al., reported that R30Q is a female-specific protective factor in pediatric CD [37]. A larger gender-stratified meta-analysis with 4707 CD patients and 4973 controls showed an association of 30Q with decreased risk of CD in females [35]. These observations raise very interesting questions including what is the "gender factor" and how does it affect IBD development? Whether the gender factor inhibits a protective effect associated with the X chromosome, or conversely, facilitates a disease promoting gene that is associated with the Y chromosome, or whether endocrine issues play a role, are currently unknown.

Acknowledgements

The authors thank Rainer Vogler, Catharina Fuerstenau, Birthe Fedders, Tanja Wesse, Tanja Henke and Lena Bossen for expert technical help with SNPlex™

genotyping, Tony Lin for his assistant with RFLP genotyping, and John Hegarty and Gaylene Webber for their help in manuscript preparation. The authors gratefully acknowledge the Gift of Life Donor Program (Philadelphia, PA) and the generosity of the organ donor families for allowing these organs that are not suitable for transplantation to be utilized to advance the understanding of human disease. This work is supported by a grant from the Philadelphia Health Care Trust (to W.A Koltun), the German Ministry of Education and Research (BMBF) through the National Genome Research Network (NGFN) and the PopGen biobank, and infrastructure support through the DFG excellence cluster "Inflammation at Interfaces" (to S. Schreiber), a research grant from Pennsylvania Department of Health Tobacco Settlement Fund (to C. MacNeill), a research grant from the Barsumian Trust, Pennsylvania State University College of Medicine (to N.J. Thomas), and a research fellowship from Germany Alexander von Humboldt Foundation (to Z. Lin).

References

[1] R.K. Russell, E.R. Nimmo and J. Satsangi, Molecular genetics of Crohn's disease, *Curr Opin Genet Dev* **14**(3) (2004), 264–270.

[2] R.K. Russell, D.C. Wilson and J. Satsangi, Unravelling the complex genetics of inflammatory bowel disease, *Arch Dis Child* **89**(7) (2004), 598–603.

[3] S. Schreiber, P. Rosenstiel, M. Albrecht et al., Genetics of Crohn disease, an archetypal inflammatory barrier disease, *Nature Reviews Genetics* **6**(5) (2005), 376–388.

[4] M. Orholm, V. Binder, T.I. Sorensen et al., Concordance of inflammatory bowel disease among Danish twins. Results of a nationwide study, *Scandinavian Journal of Gastroenterology* **35**(10) (2000), 1075–1081.

[5] C. Tysk, E. Lindberg, G. Jarnerot et al., Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking, *Gut* **29**(7) (1988), 990-996.

[6] M. Peeters, H. Nevens, F. Baert et al., Familial aggregation in Crohn's disease: increased age-adjusted risk and concordance in clinical characteristics.[see comment], *Gastroenterology* **111**(3) (1996), 597–603.

[7] J. Satsangi, C. Grootscholten, H. Holt et al., Clinical patterns of familial inflammatory bowel disease, *Gut* **38**(5) (1996), 738–741.

[8] C.O. Elson, Y. Cong, V.J. McCracken et al., Experimental models of inflammatory bowel disease reveal innate, adaptive, and regulatory mechanisms of host dialogue with the microbiota, *Immunol Rev* **206** (2005), 260–276.

[9] J.P. Hugot, M. Chamaillard, H. Zouali et al., Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease, *Nature* **411**(6837) (2001), 599–603.

[10] Y. Ogura, D.K. Bonen, N. Inohara et al., A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease, *Nature* **411**(6837) (2001), 603–606.

[11] J.D. Rioux, M.J. Daly, M.S. Silverberg et al., Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease, *Nat Genet* **29**(2) (2001), 223–228.

[12] M. Stoll, B. Corneliussen, C.M. Costello et al., Genetic variation in DLG5 is associated with inflammatory bowel disease, *Nat Genet* **36**(5) (2004), 476–480.

[13] V.D. Peltekova, R.F. Wintle, L.A. Rubin et al., Functional variants of OCTN cation transporter genes are associated with Crohn disease, *Nat Genet* **36**(5) (2004), 471-475.

[14] K. Yamazaki, D. McGovern, J. Ragoussis et al., Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease, *Human Molecular Genetics* **14**(22) (2005), 3499–3506.

[15] R.H. Duerr, K.D. Taylor, S.R. Brant, J.D. Rioux, M.J. Silverberg et al., A genome-wide association study identifies IL23R as an inflammatory bowel disease gene, *Sciencexpress* **26** (October 2006), 1–4.

[16] J. Hampe, A. Franke, P. Rosenstiel et al., A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1, *Nature Genetics* **39**(2) (2007), 207–211.

[17] J.D. Rioux, R.J. Xavier, K.D. Taylor et al., Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis, *Nature Genetics* **39**(5) (2007), 596–604.

[18] C. Libioulle, E. Louis, S. Hansoul et al., Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4, *PLoS Genetics* **3**(4) (2007), e58.

[19] Wellcome Trust Case Control C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.[see comment], *Nature* **447**(7145) (2007), 661–678.

[20] M. Parkes, J.C. Barrett, N.J. Prescott et al., Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility.[see comment], *Nature Genetics* **39**(7) (2007), 830–832.

[21] L. Gonzalez-Mariscal, A. Betanzos and A. Avila-Flores, MAGUK proteins: structure and role in the tight junction, *Semin Cell Dev Biol* **11**(4) (2000), 315–224.

[22] H. Nakamura, T. Sudo, H. Tsuiki et al., Identification of a novel human homolog of the Drosophila dlg, P-dlg, specifically expressed in the gland tissues and interacting with p55, *FEBS Letters* **433**(1-2) (1998), 63–67.

[23] S. Purmonen, T.M. Ahola, P. Pennanen et al., HDLG5/KIAA 0583, encoding a MAGUK-family protein, is a primary progesterone target gene in breast cancer cells, *International Journal of Cancer* **102**(1) (2002), 1–6.

[24] M.J. Daly, A.V. Pearce, L. Farwell et al., Association of DLG5 R30Q variant with inflammatory bowel disease.[see comment], *European Journal of Human Genetics* **13**(7) (2005), 835–839.

[25] W.G. Newman, X. Gu, R.F. Wintle et al., DLG5 variants contribute to Crohn disease risk in a Canadian population, *Human Mutation* **27**(4) (2006), 353–358.

[26] F. Friedrichs, S. Brescianini, V. Annese et al., Evidence of transmission ratio distortion of DLG5 R30Q variant in general and implication of an association with Crohn disease in men, *Human Genetics* **119**(3) (2006), 305–311.

[27] F. Friedrichs and M. Stoll, Role of discs large homolog 5, *World Journal of Gastroenterology* **12**(23) (2006), 3651–3656.

[28] S. Vermeire, M. Pierik, T. Hlavaty et al., Association of organic cation transporter risk haplotype with perianal penetrating

Crohn's disease but not with susceptibility to IBD, *Gastroenterology* **129**(6) (2005), 1845–1853.
[29] P.L. Lakatos, S. Fischer, K. Claes et al., DLG5 R30Q is not associated with IBD in Hungarian IBD patients but predicts clinical response to steroids in Crohn's disease, *Inflammatory Bowel Diseases* **12**(5) (2006), 362–368.
[30] A. Ferraris, B. Torres, D. Knafelz et al., Relationship between CARD15, SLC22A4/5, and DLG5 polymorphisms and early-onset inflammatory bowel diseases: an Italian multicentric study, *Inflammatory Bowel Diseases* **12**(5) (2006), 355–361.
[31] C. Buning, L. Geerdts, T. Fiedler et al., DLG5 variants in inflammatory bowel disease, *American Journal of Gastroenterology* **101**(4) (2006), 786–792.
[32] H.P. Torok, J. Glas, L. Tonenchi et al., Polymorphisms in the DLG5 and OCTN cation transporter genes in Crohn's disease.[see comment], *Gut* **54**(10) (2005), 1421–1427.
[33] M. Tremelling, S. Waller, F. Bredin et al., Genetic variants in TNF-alpha but not DLG5 are associated with inflammatory bowel disease in a large United Kingdom cohort, *Inflammatory Bowel Diseases* **12**(3) (2006), 178–184.
[34] C.L. Noble, E.R. Nimmo, H. Drummond et al., The contribution of OCTN1/2 variants within the IBD5 locus to disease susceptibility and severity in Crohn's disease, *Gastroenterology* **129**(6) (2005), 1854–1864.
[35] B.L. Browning, M.L. Barclay, S.A. Bingham et al., Gender-stratified analysis of DLG5 R30Q in 4707 Crohn's disease patients and 4973 controls from 12 Caucasian cohorts, JMG Online first, published on august 13, 2007 as 101136/jmg2007050773 2007.
[36] B.L. Browning, C. Huebner, I. Petermann et al., Association of DLG5 variants with inflammatory bowel disease in the New Zealand Caucasian population and meta-analysis of the DLG5 R30Q variant, *Inflammatory Bowel Diseases* **13**(9) (2007), 1069–1076.
[37] V. Biank, F. Friedrichs, U. Babusukumar et al., DLG5 R30Q variant is a female-specific protective factor in pediatric onset Crohn's disease, *American Journal of Gastroenterology* **102**(2) (2007), 391–398.
[38] H.M. Oh, J.M. Oh, S.C. Choi et al., An efficient method for the rapid establishment of Epstein-Barr virus immortalization of human B lymphocytes. [see comment], *Cell Proliferation* **36**(4) (2003), 191–197.
[39] Z. Lin, X. Cui and H. Li, Multiplex genotype determination at a large number of gene loci, *Proc Natl Acad Sci USA* **93**(6) (1996), 2582–2587.
[40] F.M. De la Vega, K.D. Lazaruk, M.D. Rhodes et al., Assessment of two flexible and compatible SNP genotyping platforms: TaqMan SNP Genotyping Assays and the SNPlex Genotyping System, *Mutation Research* **573**(1–2) (2005), 111–135.
[41] A.R. Tobler, S. Short, M.R. Andersen et al., The SNPlex genotyping system: a flexible and scalable platform for SNP genotyping, *Journal of Biomolecular Techniques: JBT* **16**(4) (2005), 398–406.
[42] wa.K. Zhang and A. Walter, Genetic factors in the etiology and potential management of inflammatory bowel disease, *Seminars in colon and rectal Surgery* **12**(1) (2001), 2–8.
[43] W. Koltun, W. Zhang, J. Thompson, M. Boyer, E. Galka and L. Poritz, Increased presence of NOD2 mutations in patients from a familial inflammatory bowel disease registry, *Tripartite Colorectal Meeting*, dublin, Ireland, 2005.
[44] S.A. Fisher, J. Hampe, A.J. Macpherson et al., Sex stratification of an inflammatory bowel disease genome search shows male-specific linkage to the HLA region of chromosome 6, *European Journal of Human Genetics* **10**(4) (2002), 259–265.