A haplotype of prostaglandin synthase 2/ cyclooxygenase 2 is involved in the susceptibility to inflammatory bowel disease

David G Cox, J Bart A Crusius, Petra HM Peeters, H Bas Bueno-de-Mesquita, A Salvador Peña, Federico Canzian

Abstract

AIM: Prostaglandin G/H synthase 2 (PTGS2 or COX2) is one of the key factors in the cellular response to inflammation. PTGS2 is expressed in the affected intestinal segments of patients with inflammatory bowel diseases (IBD). In IBD patients, non-steroidal anti-inflammatory drugs, which have been shown to reduce both the production and activity of PTGS2, may activate IBD and aggravate the symptoms. We aimed at examining genetic variants of PTGS2 that may be risk factors for IBD.

METHODS: We genotyped 291 individuals diagnosed with IBD and 367 controls from the Dutch population for the five most frequent polymorphisms of the PTGS2 gene. Clinical data were collected on all patients. DNA was extracted via normal laboratory methods. Genotyping was carried out using multiplex PCR followed by the Invader Assay and the 5’ exonuclease assay (TaqMan). New polymorphism screening was performed by pre-screening with denaturing high-performance liquid chromatography, followed by fluorescent sequencing.

RESULTS: Allele 5209G was weakly associated with Crohn’s disease (odds ratio [OR] 1.63, 95% confidence interval [CI] 1.03-2.57), and allele 8473T with ulcerative colitis (OR 1.50, 95% CI 1.00-2.27). The haplotype including both alleles showed a strong association with IBD (OR 13.15, 95% CI 3.17-116.15). This haplotype, while rare (-0.3%) in the general population, is found more frequently in patients (3.5%).

CONCLUSION: Our data suggest that this haplotype of PTGS2 contributes to the susceptibility of IBD.

© 2005 The WJG Press and Elsevier Inc. All rights reserved.

Key words: Inflammatory bowel disease; Prostaglandin G/H synthase; Cyclooxygenase; SNP; Haplotype

INTRODUCTION

Risk factors for inflammatory bowel diseases (IBD) include several environmental and genetic exposures. Diet, tobacco smoking, and childhood diseases or poor hygiene have been identified as risk factors for IBD, with some important differences between ulcerative colitis (UC) and Crohn’s disease (CD). Caucasians have been shown to be at higher risk than non-Caucasians, with North-South gradients seen in Europe.

Familial aggregation of IBD suggests a genetic component in the susceptibility to IBD. One major locus of susceptibility to IBD has been mapped by linkage analysis to chromosome 16, at the CARD15/NOD2 gene, which has a role in inflammatory responses, through activation of nuclear factor NF-κB. Two more susceptibility genes have recently been identified. Population-based association studies focused on polymorphisms of inflammatory genes, such as genes of the interleukin-1β pathway. Few showed significant associations with some degree of contradiction between studies.

Key genes involved in the regulation of the inflammatory processes, such as prostaglandin G/H synthase/cyclooxygenase (PTGS/COX), are obvious candidates to look for variants predisposing to IBD. One of the two PTGS isoforms, PTGS2/COX2, is expressed in epithelial cells and mononuclear cells in IBD, and it is induced in response to pro-inflammatory cytokines, including interleukin-1β. PTGS2 is the rate-
limiting enzyme in the production of prostaglandins. Prostaglandins are thought to be essential in the process of wound healing in the gastrointestinal tract. The use of non-steroidal anti-inflammatory drugs (NSAIDs), which inhibit both the transcription and activity of PTGS2, exacerbates the symptoms in UC and may even activate quiescent IBD. Thus, the expression of PTGS2 in the inflamed intestine might be a protective response within the wound-healing process. Consequently, polymorphisms that change the amount of prostaglandins produced in inflamed cells could cause susceptibility to IBD.

To address this hypothesis, we have studied 145 UC patients, 146 CD patients, and 367 controls from the Dutch Caucasian population, which has a high incidence of IBD. The present study was to search for associations between IBD and common single nucleotide polymorphisms (SNPs) in the PTGS2 gene.

**MATERIALS AND METHODS**

**Samples**

The recruitment of cases took place at the Department of Gastroenterology of the VU University Medical Center (VUmc), a referral center for IBD, between April 1992 and September 2000. Patients with indeterminate colitis were excluded from this study. Seventy-five UC patients were subdivided into one group with their colon in situ (116 cases), and a second group in which (procto)colectomy had to be performed during the course of disease (29 cases), indicative of severe or intractable disease (i.e., unresponsive to medical therapy). No patient was operated on prophylactic indication.

In CD, patients were subdivided according to the Vienna classification. This classification subdivides patients with CD according to age at onset (<40 years [125 cases] or ≥40 years [21 cases]), disease behavior (non-stricturing, non-penetrating [51 cases], strictureting [63 cases], penetrating [32 cases]) and location of disease (terminal ileum [50 cases], colon [32 cases], ileocolonic [62 cases], or upper gastrointestinal tract [2 cases]). Anatomical classification is defined as the maximal extent of the disease prior to the first surgical procedure, while behavior is assessed at any time during the course of the disease.

**SNP selection and genotyping**

SNPs were selected from our previously collected data to include those SNPs with a prevalence of greater than 5% in Caucasians. Among the SNPs we studied, the only one with a proven functional role is a SNP located in the promoter of PTGS2, at position 926 of GenBank entry D28235 (dbSNP rs20417), which has been shown to affect the expression of PTGS2 mRNA.

The order of samples of patients and controls was randomized in the PCR plates, so that a uniform number of patients and controls could be analyzed simultaneously in each run. EPIC samples were genotyped separately.

Three SNPs were genotyped with the Invader Assay. PCR for the Invader assay (Third Wave Technologies, Madison, WI, USA) was carried out in 25 μL reactions, using the following concentrations: 20 ng DNA, 20 μmol/L dNTPs, 1×Taq Platinum buffer, 1.5 mmol/L MgCl2, 1.25 U Taq Platinum polymerase (Life Technologies, Inc., Gaithersburg, MD, USA), 0.5 μmol/L of each primer. The primers used were as follows (nomenclature of SNPs refers to base numbers in GenBank entry D28235): PTGS2.401 (rs689465).

401F: AAC GAC TTA GGA CAT AAC TGA ATT TTC
401R: ATG GGT AGT GCT CAG GGA GGA G
PTGS2.3050 (rs5277)
3050F: CGT TGT GAA TAA CAT TCC CTT
3050R: ATT TTT CTG TAA GGA GCC TAA AA
PTGS2.5209 (rs20432)
5209F: ATG ATG TAT GCC ACA ATC TGG CTG
5209R: TTG TCT GGA ACA ACT GCT CAT CAC.

This multiplex reaction was then carried out in a Tetrad DNA Engine PCR machine (MJ Research, Waltham, MA, USA), with the following cycling conditions: 96 °C for 5 min, then 30 cycles of 96 °C for 30 s, 50 °C for 30 s, and 72 °C for 10 s, with a final extension of 72 °C for 5 min. PCR volumes were then brought up to 150 μL, and Invader reactions were carried out as per manufacturer’s instructions. Plates were read in a fluorometer with excitation and emission spectra as recommended by Third Wave.

The SNPs PTGS2.926 and PTGS2.8473 were tested using the TaqMan assay with MGB chemistry (Applied Biosystems,
Foster City, CA, USA). The assay was carried out in a 10 μL reaction, with 20 ng DNA, 1×TaqMan master mix (Applied Biosystems), 0.1 μmol/L of each primer, and 0.2 μmol/L of each probe. Primer and probe sequences are as follows: PTGS2.926 (rs20417)

926F: TTA ACT ATT TAC AGG GTA ACT GCT TAG G
926R: CTT CAC CCC CTC CTT GTT TC
926VIC: CCT TTC CCG CCT CT
926FAM: CTT TCC CCC CTC TC
PTGS2.8473 (rs5275)
8473F: ATG CAC TGA TCT CGT TTT TTG TTG TT G
8473R: GTT TCC AAC GTT ATC TTG TCC ATG A
8473VIC: TGA CAG AAA AAT AAC CAA AA
8473FAM: TGA CAG AAA AAT AAC CAA AA

The cycling conditions for these PCR reactions were as follows: 95 °C for 10 min, then 35 cycles of 95 °C for 15 s, and 60 °C (58 °C for PTGS2.926) for 60 s. Plates were then read in the ABI 7900HT sequence detection system.

**Data analysis**

For genotype data from the Invader assay, raw intensity counts were corrected with a negative control, and the ratio between the two signals (one color for each allele) was used to call each genotype. For the TaqMan assay, groups of genotypes (homozygote common, heterozygote, and homozygote rare) were determined manually within the SDS software (Applied Biosystems). Each SNP was tested in the control group to ensure that it does not deviate from Hardy-Weinberg equilibrium. Linkage disequilibrium tests were performed using macros developed by the authors in conjunction with the PHASE program[30] to reconstruct haplotype frequencies. Odds ratios (OR) and 95% confidence intervals (CI) were calculated in Stata 7.0, using logistic regression correcting for sex and age in analyzing genotypes of haplotypes in the population. All P values reported are two sided. SNPs were tested for effects on splicing using the Delila Server (http://www.lecb.ncifcrf.gov/~toms/delilaserver.html) at the Laboratory of Experimental and Computational Biology of the National Cancer Institute[31].

**New polymorphism screening**

Individuals termed “at risk” based on association testing (i.e., the 19 cases carrying the haplotype, which includes the two alleles showing increase in risk at SNPs 5 209 and 8 473) were screened for novel polymorphisms using denaturing high-performance liquid chromatography (DHPLC) via the WAVE system (Transgenomic, Omaha, NE, USA). The areas studied were the promoter region, exons 6 and 10, and the 3' UTR. The protocols followed were as published previously[27].

**RESULTS**

Figure 1 shows a graphical representation of the structure of the PTGS2 gene. The five major SNPs as well as the polymorphisms discovered as part of this study are depicted. The allele frequencies of the five major SNPs in the control group were 11.4% (PTGS2.401), 13.4% (PTGS2.926), 14.6% (PTGS2.3050), 13.5% (PTGS2.5209), and 33.9% (PTGS2.8473). All SNPs, except PTGS2.3050, were in Hardy-Weinberg equilibrium in the control group. Table 1 shows the genotype counts of the SNPs studied, among cases and controls. The SNP at position 5 209 shows an increase in the carriers of the G allele in cases as compared to controls, with a weak but significant association was found only in CD (OR = 1.63, 95%CI 1.03-2.57, P = 0.04). The SNP at position 8 473 shows an increase in the frequency of the T/T homozygote genotype in cases as compared to controls, with an association of borderline significance found only in UC (OR = 1.50, 95%CI 1.00-2.27, P = 0.05). No association was seen with the SNPs at position 401, 926, or 3 050, although it should be noted that there is a lack of homozygous rare genotypes at bp 3 050 in the control group, which is the reason for its departure from Hardy-Weinberg equilibrium.

We analyzed differences in genotype counts among the different clinical variables for CD and UC. No statistically significant differences between classification levels were observed (data not shown).

Significant levels of linkage disequilibrium were seen among the most common SNPs in the PTGS2 gene (data not shown). Table 2 shows the frequencies of haplotypes of the PTGS2 gene in both IBD patients and controls. The SNP at bp 3 050 was not included in haplotype analysis, as it was not in Hardy-Weinberg equilibrium in the control population, and this has been shown to affect the accuracy of haplotype reconstruction methods[32]. All individuals were genotyped for the SNP at bp 3 050, eliminating the risk that its lack of Hardy-Weinberg equilibrium is caused by bias due to the genotyping technique missing preferentially one genotype. Haplotype frequencies were calculated using all four valid SNPs, and all data are shown. There were no differences in haplotype frequency between the two groups of IBD patients and the controls, except for haplotype AGGT, which shows an OR of 11.9 (95%CI 2.83-105.76, P<0.00005, for UC and CD combined). Interestingly, this haplotype includes the two alleles showing an increase in risk at SNPs 5 209 and 8 473. If we consider the haplotype including the 5209.G and 8473.T alleles (by lumping haplotypes

![Figure 1](image.URL)
AGGT and GCGT), we observe an increase in its frequency in patients with IBD (0.035, UC and CD combined) as
compared to controls (0.003). This increase in frequency yields an OR of 13.15 (95%CI 3.17-116.15, 
\( P < 0.00005 \), for UC and CD combined).

In order to test the possibility that these positive associations are in reality indicative of linkage disequilibrium with yet undiscovered nearby polymorphisms, we studied by DHPLC the 19 cases who could carry the risk haplotype.

Regions of \( \text{PTGS2} \) surrounding the three SNPs, and in addition the promoter region, exon 1 and the UTRs were analyzed. No novel polymorphisms were discovered in the coding sequence of the gene. One SNP was discovered at bp 9850 (A-G), with a minor allele frequency of lower than 2%. Therefore, we considered it unlikely that polymorphisms other than 5209 or 8473 could explain the observed associations.

**DISCUSSION**

\( \text{PTGS2} \) is a critical enzyme involved in the production of prostaglandins, which are essential in the process of healing bowel wounds. The use of NSAIDs, which inhibit both the translation and activity of \( \text{PTGS2} \), can induce a flare-up\(^{[21]} \) and actually exacerbate the symptoms of IBD\(^{[20]} \). We hypothesized that polymorphisms in \( \text{PTGS2} \) might influence prostaglandin production in inflamed cells, thus affecting susceptibility to IBD. In this study, we have found a slightly increased risk of IBD associated to \( \text{PTGS2} \).5209 and \( \text{PTGS2} \).8473, and a strong association with a haplotype including alleles of these two SNPs.

\( \text{PTGS2}.5209 \) is in intron 5 of \( \text{PTGS2} \). In \textit{silico} analysis

### Table 1 Genotype frequencies and OR for UC and CD of PTGS2 SNPs

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Controls</th>
<th>UC</th>
<th>CD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n^1 )</td>
<td>%</td>
<td>( n^1 )</td>
</tr>
<tr>
<td>( \text{PTGS2.401} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>277</td>
<td>79</td>
<td>115</td>
</tr>
<tr>
<td>A/G</td>
<td>66</td>
<td>19</td>
<td>24</td>
</tr>
<tr>
<td>G/G</td>
<td>7</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>350</td>
<td>100</td>
<td>140</td>
</tr>
<tr>
<td>( \text{PTGS2.926} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td>256</td>
<td>75</td>
<td>112</td>
</tr>
<tr>
<td>G/C</td>
<td>75</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>C/C</td>
<td>8</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>339</td>
<td>100</td>
<td>139</td>
</tr>
<tr>
<td>( \text{PTGS2.3050} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td>250</td>
<td>71</td>
<td>109</td>
</tr>
<tr>
<td>G/C</td>
<td>99</td>
<td>28</td>
<td>32</td>
</tr>
<tr>
<td>C/C</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>350</td>
<td>100</td>
<td>145</td>
</tr>
<tr>
<td>( \text{PTGS2.5209} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/T</td>
<td>265</td>
<td>75</td>
<td>110</td>
</tr>
<tr>
<td>T/G</td>
<td>81</td>
<td>23</td>
<td>30</td>
</tr>
<tr>
<td>G/G</td>
<td>7</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>353</td>
<td>100</td>
<td>145</td>
</tr>
<tr>
<td>( \text{PTGS2.8473} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C/C</td>
<td>44</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>T/C</td>
<td>147</td>
<td>42</td>
<td>49</td>
</tr>
<tr>
<td>T/T</td>
<td>155</td>
<td>45</td>
<td>78</td>
</tr>
<tr>
<td>Total</td>
<td>346</td>
<td>100</td>
<td>145</td>
</tr>
</tbody>
</table>

\(^1\)Numbers may not sum up to the totals of controls or cases due to genotyping failure. All samples that did not give a reliable result in the first round of genotyping were resubmitted to up to three additional rounds of genotyping. Data points that were still not filled up after this procedure were left blank. \(^2\)All ORs are adjusted for sex and age. \(^3\)Heterozygotes and homozygotes for the rare allele have been grouped in order to improve statistical power.

### Table 2 PTGS2 haplotype frequencies and OR for IBD

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Controls</th>
<th>UC cases</th>
<th>CD cases</th>
<th>UC+CD cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n^1 )</td>
<td>( n^1 )</td>
<td>( \text{OR (95%CI)}^2 )</td>
<td>( n^1 )</td>
</tr>
<tr>
<td>AGTT</td>
<td>471 (64.7)</td>
<td>193 (65.6)</td>
<td>1.00 (ref.)</td>
<td>183 (62.7)</td>
</tr>
<tr>
<td>AGTC</td>
<td>147 (20.2)</td>
<td>55 (18.7)</td>
<td>0.91 (0.64-1.30)</td>
<td>54 (18.5)</td>
</tr>
<tr>
<td>GGCC</td>
<td>64 (8.8)</td>
<td>22 (7.5)</td>
<td>0.84 (0.50-1.39)</td>
<td>27 (9.2)</td>
</tr>
<tr>
<td>AGCG</td>
<td>25 (3.4)</td>
<td>8 (2.7)</td>
<td>0.78 (0.35-1.73)</td>
<td>7 (2.4)</td>
</tr>
<tr>
<td>AGGT</td>
<td>2 (0.3)</td>
<td>10 (3.4)</td>
<td>12.2 (2.97-inf.)</td>
<td>9 (3.1)</td>
</tr>
<tr>
<td>Rare(^4)</td>
<td>19 (2.6)</td>
<td>6 (2.0)</td>
<td>0.77 (0.31-1.90)</td>
<td>12 (4.1)</td>
</tr>
</tbody>
</table>

\(^1\)The order of SNPs in the haplotypes is \( \text{PTGS2.401}, \text{PTGS2.926}, \text{PTGS2.5209}, \) and \( \text{PTGS2.8473} \). \(^2\)UC and CD combined. \(^3\)ORs are not corrected as they are based on population level haplotype frequencies and are relative to the most common haplotype (AGTT). \(^4\)Rare haplotypes (frequency \( \leq 1\% \) in both cases and controls).
of this SNP for splice site mutations\(^{31}\) reveals that the T-G substitution creates a new splicing acceptor sequence with nearly the same strength as that found in normal splice acceptors. While this is most probably not enough of a change to exclude completely the normal splicing of exons 5-6, it could cause a “leaking”, reducing the amount of normal PTGS2 mRNA. PTGS2.8473, being in the 3' untranslated region (UTR), could affect the stability of PTGS2 mRNA\(^{18}\). In fact, the unusually long 3’ UTR plays an important role in determining the half-life of the mRNA. Various proteins bind to it and cause either acceleration or protection of the degradation of the mRNA\(^{34,35}\).

The associations we have found between CD and PTGS2.5209 G/G, and UC and PTGS2.8473 T/T are statistically significant, but of moderate importance. We propose that it is the combination of the G allele at PTGS2.5209 with the T allele at PTGS2.8473, on the same chromosome (GT haplotype), that would have the most influence on disease status.

The association between IBD and the GT haplotype could be due to linkage disequilibrium with another polymorphism that lies nearby. To test this hypothesis, we have analyzed subjects who carry the GT haplotype for the presence of novel functional polymorphisms of PTGS2, but this search has not yielded any additional candidates. Linkage disequilibrium with polymorphisms in neighboring genes is not a likely explanation either, as there is no gene mapping of at less than 100 kbp on either side of PTGS2.

We are inclined to think that this association is not a finding by chance. First of all, we have found a higher level of linkage disequilibrium between alleles at bp 5 209 and 8 473 in controls than in cases. This could mean that natural selection has exerted pressure toward the disappearance of the GT haplotype, i.e. this haplotype might negatively affect the fitness of individuals carrying it. Additionally, we have observed that this haplotype is virtually absent not only in the Dutch population, but in other Northern European populations as well (Cox et al., unpublished data).

These results are the first linking of PTGS2 to IBD. While the GT haplotype of PTGS2 is rare, the two alleles that compose it are much more frequent, they both show a modest association with the disease, and could therefore be important for a large proportion of the population. Future steps are to replicate these findings by independent studies, to determine experimentally the extent to which PTGS2 polymorphisms alter the expression or function of the gene, and to discover genetic or environmental factors that may alter the risk of IBD interacting with PTGS2 alleles.

ACKNOWLEDGMENTS
The authors thank Pietro Ferrari (Nutrition and Cancer Group, IARC) for discussion on statistical issues, Stéphanie Monnier (Genome Analysis Team, IARC) for bioinformatic support and AA van Bodegraven MD, PhD for facilitating the clinical database of the IBD patients at the VUmc.

REFERENCES


27 Cox D, Boillot C, Canzian F. Data mining: Efficiency of using sequence databases for polymorphism discovery. *Hum Mutat* 2001; 17: 141-150


