[1] Discipline of Nutrition, University of Auckland, Auckland, New Zealand; [2] Department of Statistics, University of Auckland, Auckland, New Zealand; [3] Nutrigenomics, New Zealand; [4] Department of Gastroenterology, Christchurch Hospital, Christchurch, New Zealand; [5] Department of Medicine, Christchurch School of Medicine, Christchurch, New Zealand; [6] EPIC Norfolk and MRC Centre for Nutrition and Cancer Prevention and Survival, University of Cambridge, Cambridge, UK; [7] Department of Medicine II – Grosshadern, University of Munich, Munich, Germany; [8] Department of Gastroenterology, Hepatology and Endocrinology, Charité, Campus Mitte, Universitätsmedizin Berlin, Berlin, Germany; [9] Gastroenterology Unit, IRCCS, Bambino Gesù Children's Hospital, Rome, Italy; [10] Department of Pediatrics, University of Rome La Sapienza, Rome, Italy; [11] IRCCS CSS-Mendel Institute, Rome, Italy; [12] Gastrointestinal Unit, University of Edinburgh, Western General Hospital, Edinburgh, UK; [13] Department of Medical and Molecular Genetics, King's College London School of Medicine, London, UK; [14] Department of Restorative Dentistry and Periodontology, University of Munich, Munich, Germany; [15] Department of Gastroenterology, UZ Gasthuisberg, Leuven, Belgium; [16] First Department of Medicine, Csolnoky F. County Hospital, Veszprem, Hungary; [17] First Department of Medicine, Semmelweis University, Budapest, Hungary; [18] Department of Gastroenterology, IRCCS, Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy; [19] Department of Medicine, University of Toronto, Toronto, Ontario, Canada; [20] Department of Immunology, University of Toronto, Toronto, Ontario, Canada; [21] Department of Medical Genetics and Microbiology, University of Toronto, Toronto, Ontario, Canada; [22] Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada; [23] Max-Planck Institute of Psychiatry, Munich, Germany; [24] Department of Medical Genetics, University of Manchester, Manchester, UK; [25] Department of Gastroenterology, Addenbrookes Hospital, Cambridge, UK; [26] Department of Obstetrics and Gynaecology, University of Auckland, Auckland, New Zealand; [27] Department of Surgery – Innenstadt, University of Munich, Munich, Germany; [28] Department of Pediatrics, Charité, Campus Virchow-Klinikum, Universitätsmedizin Berlin, Berlin, Germany; [29] Department of Gastroenterology and Hepatology, Charité, Campus Virchow-Klinikum, Universitätsmedizin Berlin, Berlin, Germany

Correspondence to:
B L Browning, Discipline of Nutrition, The University of Auckland, Private Bag 92019, Auckland, New Zealand; b.browning@auckland.ac.nz

Gender-stratified analysis of DLG5 R30Q in 4707 patients with Crohn disease and 4973 controls from 12 Caucasian cohorts

B L Browning,[1,2,3] V Annese,[18] M L Barclay,[4,5] S A Bingham,[6] S Brand,[7] C Büning,[8] M Castro,[9] S Cucchiara,[10] B Dallapiccola,[11] H Drummond,[12] L R Ferguson,[1,3] A Ferraris,[11] S A Fisher,[13] R B Gearry,[4,5] J Glas,[7,14] L Henckaerts,[15] C Huebner,[1,3] D Knafelz,[9] L Lakatos,[16] P L Lakatos,[17] A Latiano,[18] X Liu,[19,20,21,22] C Mathew,[13] B Müller-Myhsok,[23] W G Newman,[24] E R Nimmo,[12] C L Noble,[12] O Palmieri,[18] M Parkes,[25] I Petermann,[1,3] P Rutgeerts,[15] J Satsangi,[12] A N Shelling,[3,26] K A Siminovitch,[19,20,21,22] H-P Török,[7,27] M Tremelling,[25] S Vermeire,[15] M R Valvano,[18] H Witt[28,29]

ABSTRACT

Background: *DLG5* p.R30Q has been reported to be associated with Crohn disease (CD), but this association has not been replicated in most studies. A recent analysis of gender-stratified data from two case–control studies and two population cohorts found an association of *DLG5* *30Q* with increased risk of CD in men but not in women and found differences between *30Q* population frequencies for males and females. Male–female differences in population allele frequencies and male-specific risk could explain the difficulty in replicating the association with CD.

Methods: *DLG5* R30Q genotype data were collected for patients with CD and controls from 11 studies that did not include gender-stratified allele counts in their published reports and tested for male–female frequency differences in controls and for case–control frequency differences in men and in women.

Results: The data showed no male–female allele frequency differences in controls. An exact conditional test gave marginal evidence that *30Q* is associated with decreased risk of CD in women (p = 0.049, OR = 0.87, 95% CI 0.77 to 1.00). There was also a trend towards reduced *30Q* frequencies in male patients with CD compared with male controls, but this was not significant at the 0.05 level (p = 0.058, OR = 0.87, 95% CI 0.74 to 1.01). When data from this study were combined with previously published, gender-stratified data, the *30Q* allele was found to be associated with decreased risk of CD in women (p = 0.010, OR = 0.86, 95% CI 0.76 to 0.97), but not in men.

Conclusion: *DLG5* *30Q* is associated with a small reduction in risk of CD in women.

Inflammatory bowel disease (IBD; OMIM 266600) comprises Crohn disease (CD; OMIM 266600) and ulcerative colitis (UC; OMIM 191390). Epidemiological and genetic studies have demonstrated that genes play an important role in the pathogenesis of IBD.[1–6] Genetic variants associated with CD have been identified and replicated in the *CARD15/NOD2* gene (OMIM 605956)[7–9] and in the chromosome 5q31 region (OMIM 606348),[10–13] and recently, the first generation of whole-genome association studies have identified multiple new susceptibility variants.[14–19]

In 2004, Stoll *et al* described an association between IBD and *DLG5* variants (*DLG5*: Drosophila Discs Large Homolog 5 (OMIM 604090). In particular, p.R30Q (c.89G→A, rs1248696, previously reported as c.113G→A) was found to be associated with IBD in a German family cohort. Stoll *et al*[20] replicated the R30Q association with CD in a European case–control cohort and found a non-significant overtransmission of the *30Q* allele in patients with CD in a German/UK family cohort. Subsequently, Daly *et al*[21] replicated the association of *30Q* with CD in one of two case–control cohorts and in a family cohort. However, subsequent studies have failed to replicate the association of *DLG5* *30Q* with CD.[22–34]

Recently, Friedrichs *et al*[35] examined gender-stratified data from the case–control cohorts showing association in the initial two *DLG5* studies,[20 21] and found that the *DLG5* *30Q* allele was a risk factor for CD in men but not in women, and that the *30Q* allele frequency was significantly lower in male compared with female controls (0.05 vs 0.11). Friedrichs *et al*[35] genotyped R30Q in two additional control samples: a sample of 190 male and 271 female participants from Germany and in a newborn sample of 301 male and 299 female infants from Wisconsin, USA. For all three samples (combined control cohort, the German population sample and the Wisconsin population sample) the *DLG5* *30Q* allele frequency was lower in male than in female subjects and the differences were significant at the 0.05 level. The presence of the male–female allele frequency differences in the newborn sample suggests that the allele frequency differences are due to prenatal processes rather than selection later in life and Friedrichs *et al*[35] proposed that the allele frequency differences were a consequence of a gender-dependent transmission ratio distortion.

Differences in male and female allele frequencies combined with male-specific risk could explain the difficulties that have been encountered in replicating the association of *30Q* with CD, as the estimated effect size would depend on the proportion of male cases and controls in the study. If the minor allele increases risk in men but not in

women and if the frequency in male patients with CD is less than the minor allele frequency in females in a population, then the *30Q* allele would be a risk allele in a study of female patients with CD and male controls, but would be a protective allele in a study of male patients with CD and female controls. Replication of association could be further complicated by differences in effect sizes between populations, owing to interaction of R30Q with other genetic and environmental factors that vary between populations.

Most studies investigating R30Q were published prior to the report of Friedrichs *et al*,[35] which first described differences in male and female R30Q population frequencies and male–specific risk. Consequently, in most *DLG5* studies, genotype data stratified by gender have not been reported. In this study, we analysed male and female R30Q allele counts from 11 previous *DLG5* case–control studies that did not report gender-stratified data. We tested R30Q for association with CD in men and in women, and tested for differences in male and female control allele frequencies.

METHODS
Study participants
Published association studies of *DLG5* were identified through PubMed using a search with the keyword "DLG5". All association studies had sampled from Caucasian populations. The limited available evidence indicates that the *DLG5* R30Q polymorphism is absent in populations from Asia and Africa.[36 37] We identified 12 case–control association studies of *DLG5* R30Q and IBD in Caucasian populations, which had not reported gender-stratified R30Q allele counts or frequencies. One of these 12 studies had a relatively high proportion of non-Caucasian participants.[38] Gender-stratified data from the remaining 11 DLG5 case–control studies[22–27 29 30 32–34] are included in the current study and represent 4707 patients with CD and 4973 controls from 12 case–control cohorts.

Allele counts reported in the current study may be greater or less than allele counts in the original study if gender was not available for some participants or if additional individuals from the same population were genotyped for DLG5 R30Q subsequent to the initial study. Our dataset does not include data from the Ashkenazi Jewish participants in the study of Newman *et al*,[29] nor the CD and control patients in the study of Cucchiara *et al*[32] who had previously been included in the

study of Friedrich *et al*,[35]. Local ethics committee approval was obtained for each study that contributed gender-stratified data to the current study.

All mutations are numbered at the cDNA level, indicated by a "c." before the number. Position +1 corresponds to the A of the ATG translation initiation codon located at nucleotide 401 in the NM_004747.3 *DLG5* mRNA reference sequence.

Statistical analysis
Statistical analysis was performed using R V.2.2.0 software.[39] For comparison of male and female *30Q* population frequencies, allelic odds ratios (ORs) <1.0 correspond to a lower *30Q* allele frequency in male than in female subjects. For case–control analysis, allelic ORs <1.0 correspond to a lower *30Q* allele frequency in patients with CD than in controls. Allelic ORs with 95% confidence intervals (CIs) were calculated under the assumption of Hardy–Weinberg equilibrium both in cases and in controls.

The combined evidence for association was assessed with an exact conditional test (conditional upon cohort). The exact test was used in preference to the Cochran–Mantel–Haenszel test because the exact test does not use an asymptotic approximation. Heterogeneity between studies was assessed using Cochran's Q statistic,[40] which has a χ^2 distribution under the null hypothesis of no heterogeneity between studies, and using the I^2 statistic,[41] which is an estimate of the proportion of the total variation that is due to heterogeneity between studies. When there was evidence for heterogeneity between studies, we performed a meta-analysis under a random effects model[42] to estimate the mean OR. Under a random effects model, the variance of the allelic OR for each study population is assumed to be drawn from a random distribution with variance δ and the variance of the estimated allelic OR is the sum of the sampling variance and δ. Estimates of mean OR under a random effects model are calculated by weighting each study by the reciprocal of the sampling variance of the OR to accounts for differences in sample sizes.
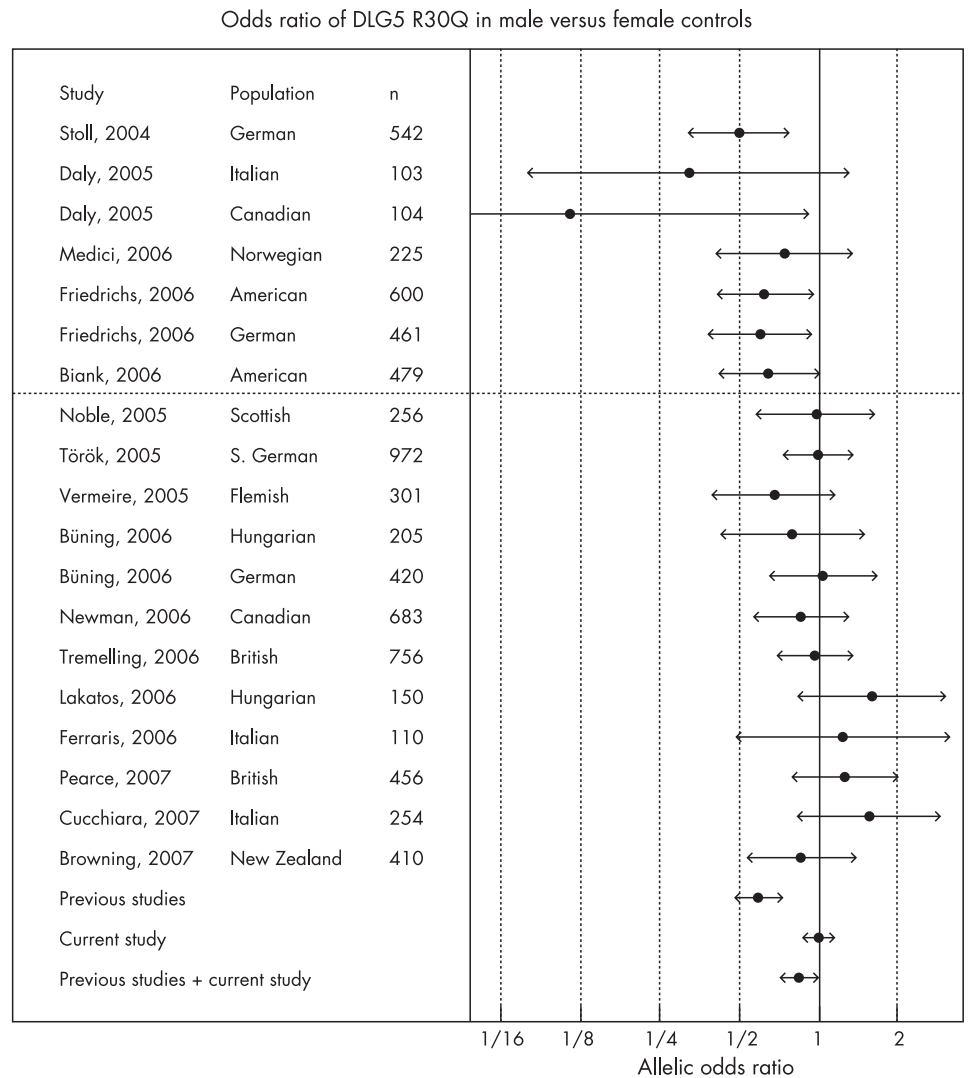
RESULTS
Male and female *DLG5* R30Q allele counts for 12 case–control cohorts are given in table 1. Genotype counts in CD cases and in controls were tested for departures from Hardy–Weinberg equilibrium using an exact test.[43] None of the tests for Hardy–

Table 1 Male and female *DLG5* R30Q allele frequency data reported in this study

Study	Population	Male patients with CD			Male controls			OR (95% CI)	Female patients with CD			Female controls			OR (95% CI)
		n	*30Q*	%	n	*30Q*	%		n	*30Q*	%	n	*30Q*	%	
Browning, 2007[34]	New Zealand	139	37	13.3	173	32	9.2	1.51 (0.91 to 2.49)	250	60	12.0	237	50	10.5	1.16 (0.78 to 1.72)
Cucchiara, 2007[32]	Italian	92	5	2.7	153	36	11.8	0.21 (0.08 to 0.54)	65	7	5.4	101	16	7.9	0.66 (0.26 to 1.66)
Pearce, 2007[33]	British	270	40	7.4	178	37	10.4	0.69 (0.43 to 1.10)	360	82	11.4	278	47	8.5	1.39 (0.95 to 2.03)
Ferraris, 2006[26]	Italian	69	14	10.1	63	13	10.3	0.98 (0.44 to 2.18)	65	9	6.9	47	8	8.5	0.80 (0.30 to 2.16)
Lakatos, 2006[27]	Hungarian	309	71	11.5	72	26	18.1	0.59 (0.36 to 0.96)	330	73	11.1	78	19	12.2	0.90 (0.52 to 1.54)
Tremelling, 2006[30]	British	168	39	11.6	330	75	11.4	1.02 (0.68 to 1.55)	326	64	9.8	426	100	11.7	0.82 (0.59 to 1.14)
Newman, 2006[29]	Canadian	141	27	9.6	171	33	9.6	0.99 (0.58 to 1.69)	172	26	7.6	512	114	11.1	0.65 (0.42 to 1.02)
Büning, 2006[25]	German	90	14	7.8	212	42	9.9	0.77 (0.41 to 1.44)	148	20	6.8	208	40	9.6	0.68 (0.39 to 1.19)
Büning, 2006[25]	Hungarian	65	11	8.5	95	19	10.0	0.83 (0.38 to 1.81)	82	12	7.3	110	27	12.3	0.56 (0.28 to 1.15)
Török, 2005[23]	South German	225	45	10.0	594	128	10.8	0.92 (0.64 to 1.32)	252	50	9.9	378	82	10.8	0.91 (0.62 to 1.31)
Noble, 2005[22]	Scottish	142	28	9.9	128	33	12.9	0.74 (0.43 to 1.26)	214	45	10.5	128	34	13.3	0.77 (0.48 to 1.23)
Vermeire, 2005[24]	Flemish	295	62	10.5	137	24	8.8	1.22 (0.75 to 2.01)	438	90	10.3	164	41	12.5	0.80 (0.54 to 1.19)
Cumulative report	—	2005	—	—	2306	—	—	0.87 (0.74 to 1.01)	2702	—	—	2667	—	—	0.87 (0.77 to 1.00)

Study, population, case and control sample sizes, *30Q* allele counts, *30Q* allele frequencies and ORs with 95% CIs are given.
The OR and 95% CI for the cumulative data were calculated using an exact conditional test.

Figure 1 Odds ratio of *DLG5* R30Q in male and female controls. Odds ratios <1.0 correspond to decreased *30Q* frequency in males. The horizontal dotted line divides previously published data (above the line) and data in the current study (below the line).

Odds ratio of DLG5 R30Q in male versus female controls

Study	Population	n
Stoll, 2004	German	542
Daly, 2005	Italian	103
Daly, 2005	Canadian	104
Medici, 2006	Norwegian	225
Friedrichs, 2006	American	600
Friedrichs, 2006	German	461
Biank, 2006	American	479
Noble, 2005	Scottish	256
Török, 2005	S. German	972
Vermeire, 2005	Flemish	301
Büning, 2006	Hungarian	205
Büning, 2006	German	420
Newman, 2006	Canadian	683
Tremelling, 2006	British	756
Lakatos, 2006	Hungarian	150
Ferraris, 2006	Italian	110
Pearce, 2007	British	456
Cucchiara, 2007	Italian	254
Browning, 2007	New Zealand	410
Previous studies		
Current study		
Previous studies + current study		

Allelic odds ratio

Weinberg equilibrium were significant at the 0.05 level after correcting for multiple testing using a Bonferroni correction.

There was no support for differences in *30Q* allele frequencies between male and female controls in our data (exact test OR = 0.99, 95% CI 0.87 to 1.13) and there was no evidence of heterogeneity of OR between populations (Q = 8.83, 11 degrees of freedom (df), p = 0.64, $I^2 = 0.00$).

In our female case–control data, there was a lower *30Q* frequency in female patients with CD compared with female controls (exact test p = 0.049, OR = 0.87, 95% CI 0.77 to 1.00) and there was no evidence of heterogeneity of OR between female populations (Q = 12.7, 11 df, p = 0.32, $I^2 = 0.13$).

In our male case–control data, there was marginal evidence of heterogeneity of OR (Q = 19.9, 11 df, p = 0.047, $I^2 = 0.45$), so we performed analysis using a random effects model, which allows for heterogeneity of ORs between populations, in addition to analysis using an exact test, which assumes homogeneity of effect between populations. For the male data, the exact test was not significant at the 0.05 level (p = 0.058, OR = 0.87, 95% CI 0.74 to 1.01) and the 95% CI for the mean OR under a random effects model included 1.0 (mean OR = 0.86, 95% CI 0.70 to 1.06).

Combining our male and female data, there was marginal evidence for heterogeneity of ORs (Q = 20.0, 11 df, p = 0.046, $I^2 = 0.45$) and our combined male and female data supported an association of *30Q* with reduced risk of CD (random effects mean OR = 0.85, 95% CI 0.74 to 0.98).

In summary, we found no male–female differences in *DLG5* R30Q frequencies among controls, but our data gives marginal evidence that the 30Q was associated with reduced risk of CD in both men and women.

We next compared our findings with previously published gender-stratified R30Q data,[28 31 35] which include gender-stratified data from Stoll *et al*[20] and Daly *et al*[21] Population, sample sizes, ORs and CIs are shown in the figures for male controls and female controls (fig 1), for female patients with CD and female controls (fig 2) and for male patients with CD and male controls (fig 3). In all three figures, the sampled population, the sample size, the sample OR and 95% CIs for the sample OR are given. The horizontal dotted line divides previously published data (above the line) and data in the current study (below the line). The ORs with 95% CI for the pooled data from previous studies, for the pooled data from the current study and for the pooled data from all studies are calculated using a random effects model. Allelic ORs are graphed using a logarithmic scale for the *x* axis.

Published gender-stratified data from previous studies show no evidence of heterogeneity of ORs in male controls versus female controls, but give strong evidence for reduced *30Q* frequency in males compared with females (exact test p = 10^{-7},

Figure 2 Odds ratio of *DLG5* R30Q in female patients with CD and female controls. Odds ratios<1.0 correspond to decreased *30Q* frequency in patients with CD. The horizontal dotted line divides previously published data (above the line) and data in the current study (below the line).
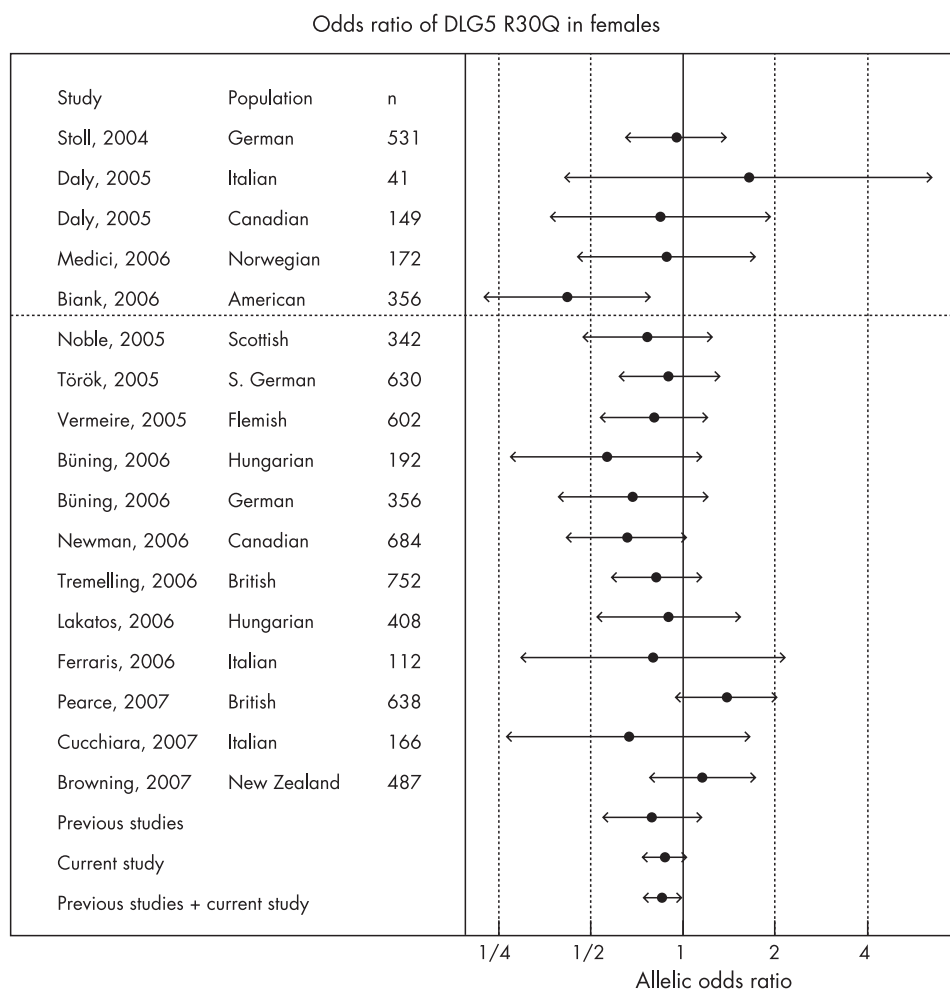
OR = 0.58, 95% CI 0.47 to 0.71).[28][31][35]. This contrasts with our data, which showed no male–female *30Q* frequency differences (OR = 0.99, see fig 1). When our data were combined with previously published data using a random effects model, the estimated mean OR was 0.84 with 95% CI 0.72 to 0.98. Our large sample of 4973 controls showed no difference in male and female *30Q* frequencies, but the magnitude of the effect described in previous studies is sufficiently large that analysis of the combined data shows allele frequency differences. The discordance between our data and previously published data suggests that data from additional studies are needed and that it would be premature to conclude that male–female allele frequency differences exist or are absent for *DLG5* R30Q.

Female data from previous studies show a trend toward reduced *30Q* allele frequency in female patients with CD compared with female controls (exact test p = 0.088, OR = 0.79, 95% CI 0.61 to 1.04). This is consistent with our data. When our data were combined with previously published female data, there was a significant reduction in 30Q frequency in female patients with CD compared with female controls (exact test p = 0.010, OR = 0.86, 95% CI 0.76 to 0.97). None of the female datasets (our female data, previously published female data or combined female data), showed evidence for heterogeneity of ORs (combined data Q = 19.0, 16 df, p = 0.27, I² = 0.16).

Published male data from previous studies show a significant increase in *30Q* frequency in male patients with CD compared with male controls (exact test p = 0.0020, OR = 1.60, 95% CI 1.18 to 2.17). In contrast, our data show a reduced 30Q

frequency in male patients with CD compared with controls (OR = 0.87). Consequently, when our male data are combined with previously published male data, there is evidence for heterogeneity of ORs between studies (Q = 36.5, 16 df, p = 0.0025, I² = 0.56), but no evidence that the mean OR differs from 1.0 under a random effects model (mean OR = 0.99, 95% CI 0.80 to 1.22).
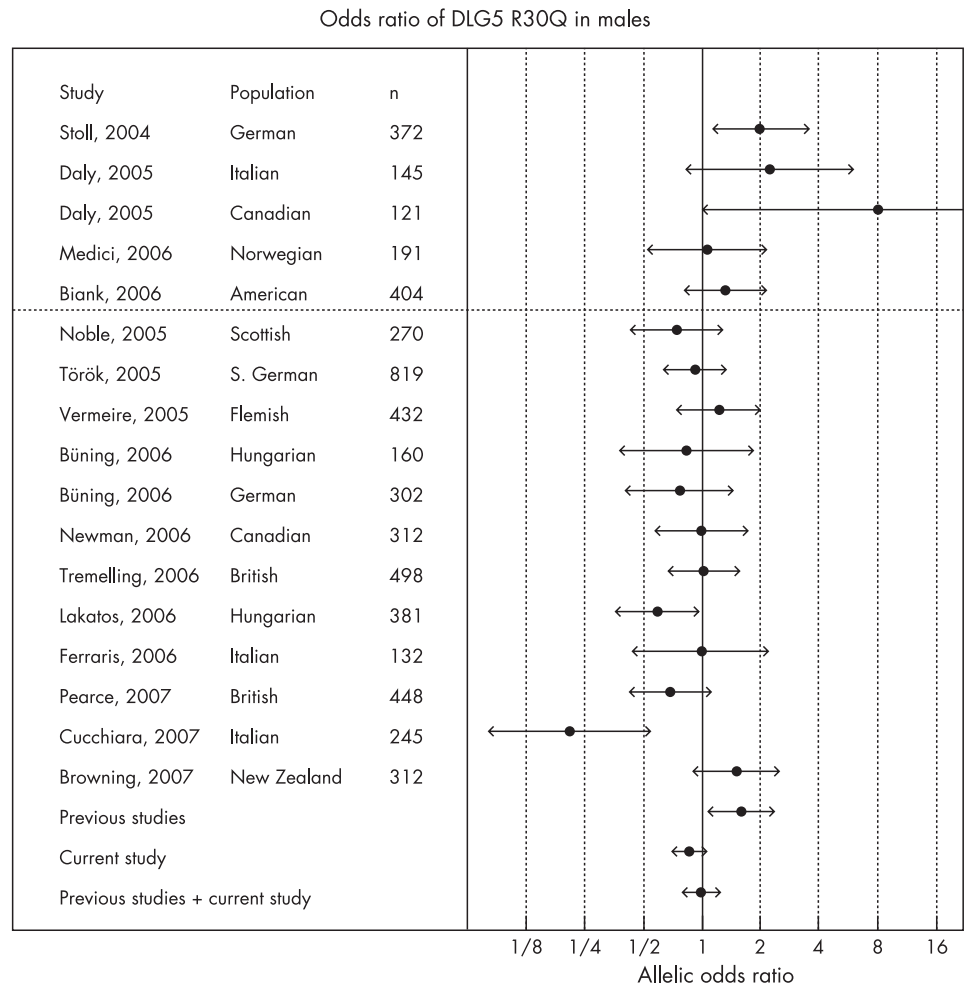
DISCUSSION

Evidence for association of the *30Q* allele with increased risk for CD is provided by the initial two *DLG5* studies[20][21] but attempts to replicate this association generally have not been successful.[22–34] The first four published whole-genome association studies for CD, three of which genotyped the R30Q variant, did not report DLG5 variants to be associated with CD in their initial published analyses.[15][16][18][19] It has been suggested that differences in male and female population allele frequencies and different effects in males compared with females could make it difficult to replicate the association of *30Q* with CD.

In this study, we report gender-stratified data for 4311 males and 5369 females from 12 case–control cohorts. We have tested our data for male–female control frequency differences and for case–control frequency differences in men and in women and we place our data in the context of previously published gender-stratified R30Q data.

The most striking difference between the data in the current study and data in previous studies is the absence of male and female population allele frequency differences for *DLG5* R30Q

Figure 3 Odds ratio of *DLG5* R30Q in male patients with CD and male controls. Odds ratios <1.0 correspond to decreased *30Q* frequency in patients with CD. The horizontal dotted line divides previously published data (above the line) and data in the current study (below the line).

Odds ratio of DLG5 R30Q in males

Study	Population	n
Stoll, 2004	German	372
Daly, 2005	Italian	145
Daly, 2005	Canadian	121
Medici, 2006	Norwegian	191
Biank, 2006	American	404
Noble, 2005	Scottish	270
Török, 2005	S. German	819
Vermeire, 2005	Flemish	432
Büning, 2006	Hungarian	160
Büning, 2006	German	302
Newman, 2006	Canadian	312
Tremelling, 2006	British	498
Lakatos, 2006	Hungarian	381
Ferraris, 2006	Italian	132
Pearce, 2007	British	448
Cucchiara, 2007	Italian	245
Browning, 2007	New Zealand	312
Previous studies		
Current study		
Previous studies + current study		

Allelic odds ratio: 1/8 1/4 1/2 1 2 4 8 16

in our data (OR = 0.99), but the presence of strong differences in previous data (OR = 0.58, p = 10^{-7}). It is unlikely that sampling variability alone can account for this difference. When we extend CIs to have 99.5% confidence, the CI for the OR in our data and the CI for the OR in previously published data are disjoint.

We considered whether population differences could account for the absence of male–female allele frequency differences in our study. The use of "healthy" controls versus unphenotyped population controls is unlikely to explain the differences as CD has <1% prevalence.[44] Different age distributions between control cohorts is also unlikely to explain the difference, as previous studies have found male–female allele frequency differences in both newborn and adult population samples. Undetected population stratification could result in the presence (or absence) of male–female allele frequency differences if male and female controls tended to be sampled from different subpopulations.

We investigated the possibility that discrepancies between our data and previously published data could be due to genotyping artefacts. In particular, we examined the possibility that nonspecific primer binding could lead to different distributions of missing and incorrect genotypes in men and women. We compared the DLG5 cDNA sequence (NM_004747) against the human genome using the NCBI MEGABLAST tool (http://www.ncbi.nlm.nih.gov/BLAST/), but found no homologies other than DLG5. A variety of genotyping platforms has been used in the published DLG5 studies and we were not able to

trace the difference between published data and our data to the use of any single genotyping platform.

For the studies that have reported the primers used to genotype the R30Q variant, we tested the specificity of primers using the NCBI Nucleotide BLAST tool ("search for short nearly exact matches"). We did not identify any hits on the X or Y chromosome that could lead to non-specific binding in men or women.

A combination of sampling variability combined with a reporting bias may help explain the difference between the set of DLG5 studies that have reported gender-stratified data and the set of DLG5 studies in the current study. Studies that observe a significant male–female allele frequency difference in controls would presumably be more likely to report gender-stratified control data.

Fortunately, many of the whole-genome association studies will genotype the *DLG5* R30Q variant in their control cohorts. The R30Q variant is not highly correlated with any other known variant,[20] and no single-nucleotide polymorphisms genotyped for the HapMap project[37] that are within 200 kb of R30Q have pairwise squared correlation coefficients $r^2 > 0.3$ in Caucasians. Consequently, the *DLG5* R30Q polymorphism will generally be genotyped in whole-genome association studies that use tagging markers, such as those using the Illumina genotyping platform.[45] We expect that these whole-genome association studies will provide the additional genotype data necessary to confirm or refute the presence of male–female population allele frequency differences for *DLG5* R30Q.

Key points

▶ We performed a large study to assess gender-specific risk of Crohn disease for *DLG5* R30Q.

▶ *DLG5* R30Q is not associated with increased risk of Crohn disease in men in our dataset.

▶ Our data show *DLG5* R30Q to be associated with a small decrease in risk of Crohn disease in women.

Previously published male data show higher *30Q* frequency in male patients with CD compared with male controls, but our male data show a trend toward lower *30Q* frequencies in male patients with CD compared with male controls. When the data are combined, the estimated mean OR is close to 1.0 (OR = 0.97), so the association of DLG5 R30Q with CD in men appears doubtful. We note that there is evidence for heterogeneity of ORs between studies in the combined data (p = 0.0020). If heterogeneity between ORs in different male populations does exist, then an estimated mean OR of 0.97 indicates that the 30Q allele is associated with a reduced risk in some male populations and increased risk in other male populations, which seems unlikely.

Our female data are consistent with previously published female data, and the combined data indicate that 30Q is associated with reduced risk of CD. Thus, although a significant decrease in female *30Q* frequency is seen in only 1 of 17 cohorts studied (fig 2), the trend toward lower ORs seen in figure 2 supports an association of *30Q* with reduced risk of CD in women (exact test p = 0.010, OR = 0.86, 95% CI 0.76 to 0.97). The presence of female-specific effects could arise out of interaction with gender-specific risk factors such as oral contraceptive use.[46 47] Interestingly, Purmonen *et al*[48] showed that DLG5 is a hormonally regulated member of the MAGUK gene family and suggested DLG5 to be a primary progesterone target gene in human breast cancer cells. This makes biological sense, as progesterone is uniquely a female hormone, produced at significant levels from puberty. Synthetic forms of progesterone are also a significant component of the combined oral contraceptive pill and a major component of the progesterone-only contraceptive pill. These associations may help provide an explanation for observed gender specific differences in incidence of CD.[49]

The estimated OR of 0.84 for the minor *30Q* allele is equivalent to an OR of 1.16 for the major *R30* allele. The modest *DLG5 30Q* allele frequency (~0.10) and the small estimated effect size in women indicates that large collaborative efforts will be needed to confirm the association of 30Q with reduced risk of CD in women.

REFERENCES

1. **Orholm M,** Munkholm P, Langholz E, Nielsen OH, Sorensen TI, Binder V. Familial occurrence of inflammatory bowel disease. *N Engl J Med* 1991;**324**:84–8.
2. **Meucci G,** Vecchi M, Torgano G, Arrigoni M, Prada A, Rocca F, Curzio M, Pera A, de Franchis R. Familial aggregation of inflammatory bowel disease in northern Italy: a multicenter study. The Gruppo di Studio per le Malattie Infiammatorie Intestinali (IBD Study Group). *Gastroenterology* 1992;**103**:514–19.
3. **Probert CS,** Jayanthi V, Hughes AO, Thompson JR, Wicks AC, Mayberry JF. Prevalence and family risk of ulcerative colitis and Crohn's disease: an epidemiological study among Europeans and south Asians in Leicestershire. *Gut* 1993;**34**:1547–51.
4. **Orholm M,** Binder V, Sorensen TI, Rasmussen LP, Kyvik KO. Concordance of inflammatory bowel disease among Danish twins. Results of a nationwide study. *Scand J Gastroenterol* 2000;**35**:1075–81.
5. **Mathew CG,** Lewis CM. Genetics of inflammatory bowel disease: progress and prospects. *Hum Mol Genet* 2004;**13**(suppl.):R161–8.
6. **van Heel DA,** Fisher SA, Kirby A, Daly MJ, Rioux JD, Lewis CM. Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. *Hum Mol Genet* 2004;**13**:763–70.
7. **Hugot JP,** Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J, Almer S, Tysk C, O'Morain CA, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel JF, Sahbatou M, Thomas G. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001;**411**:599–603.
8. **Ogura Y,** Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, Britton H, Moran T, Karaliuskas R, Duerr RH, Achkar JP, Brant SR, Bayless TM, Kirschner BS, Hanauer SB, Nunez G, Cho JH. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 2001;**411**:603–6.
9. **Economou M,** Trikalinos TA, Loizou KT, Tsianos EV, Ioannidis JP. Differential effects of NOD2 variants on Crohn's disease risk and phenotype in diverse populations: a metaanalysis. *Am J Gastroenterol* 2004;**99**:2393–404.
10. **Rioux JD,** Daly MJ, Silverberg MS, Lindblad K, Steinhart H, Cohen Z, Delmonte T, Kocher K, Miller K, Guschwan S, Kulbokas EJ, O'Leary S, Winchester E, Dewar K, Green T, Stone V, Chow C, Cohen A, Langelier D, Lapointe G, Gaudet D, Faith J, Branco N, Bull SB, McLeod RS, Griffiths AM, Bitton A, Greenberg GR, Land er ES, Siminovitch KA, Hudson TJ. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat Genet* 2001;**29**:223–8.
11. **Mirza MM,** Fisher SA, King K, Cuthbert AP, Hampe J, Sand erson J, Mansfield J, Donaldson P, Macpherson AJ, Forbes A, Schreiber S, Lewis CM, Mathew CG. Genetic evidence for interaction of the 5q31 cytokine locus and the CARD15 gene in Crohn disease. *Am J Hum Genet* 2003;**72**:1018–22.
12. **Fisher SA,** Hampe J, Onnie CM, Daly MJ, Curley C, Purcell S, Sand erson J, Mansfield J, Annese V, Forbes A, Lewis CM, Schreiber S, Rioux JD, Mathew CG. Direct or indirect association in a complex disease: the role of SLC22A4 and SLC22A5 functional variants in Crohn disease. *Hum Mutat* 2006;**27**:778–785.
13. **Silverberg MS,** Duerr RH, Brant SR, Bromfield G, Datta LW, Jani N, Kane SV, Rotter JI, Philip Schumm L, Hillary Steinhart A, Taylor KD, Yang H, Cho JH, Rioux JD, Daly MJ. Refined genomic localization and ethnic differences observed for the IBD5 association with Crohn's disease. *Eur J Hum Genet* 2007;**15**:328–35.
14. **Duerr RH,** Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, Abraham C, Regueiro M, Griffiths A, Dassopoulos T, Bitton A, Yang H, Targan S, Datta LW, Kistner EO, Schumm LP, Lee AT, Gregersen PK, Barmada MM, Rotter JI, Nicolae DL, Cho JH. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 2006;**314**:1461–3.
15. **Hampe J,** Franke A, Rosenstiel P, Till A, Teuber M, Huse K, Albrecht M, Mayr G, De La Vega FM, Briggs J, Gunther S, Prescott NJ, Onnie CM, Hasler R, Sipos B, Folsch UR, Lengauer T, Platzer M, Mathew CG, Krawczak M, Schreiber S. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 2007;**39**:207–211.
16. **Libioulle C,** Louis E, Hansoul S, Sand or C, Farnir F, Franchimont D, Vermeire S, Dewit O, de Vos M, Dixon A, Demarche B, Gut I, Heath S, Foglio M, Liang L, Laukens D, Mni M, Zelenika D, Van Gossum A, Rutgeerts P, Belaiche J, Lathrop M, Georges M. Novel crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. *PLoS Genet* 2007;**3**:e58.
17. **Parkes M,** Barrett JC, Prescott NJ, Tremelling M, Anderson CA, Fisher SA, Roberts RG, Nimmo ER, Cummings FR, Soars D, Drummond H, Lees CW, Khawaja SA, Bagnall R, Burke DA, Todhunter CE, Ahmad T, Onnie CM, McArdle W, Strachan D, Bethel G, Bryan C, Lewis CM, Deloukas P, Forbes A, Sand erson J, Jewell DP, Satsangi J, Mansfield JC, Cardon L, Mathew CG. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 2007;**39**:830–2.
18. **Rioux JD,** Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A, Green T, Kuballa P, Barmada MM, Datta LW, Shugart YY, Griffiths AM, Targan SR, Ippoliti AF, Bernard EJ, Mei L, Nicolae DL, Regueiro M, Schumm LP, Steinhart AH, Rotter JI, Duerr RH, Cho JH, Daly MJ, Brant SR. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 2007;**39**:596–604.
19. **The Wellcome Trust Case Control Consortium**. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;**447**:661–78.
20. **Stoll M,** Corneliussen B, Costello CM, Waetzig GH, Mellgard B, Koch WA, Rosenstiel P, Albrecht M, Croucher PJ, Seegert D, Nikolaus S, Hampe J, Lengauer T, Pierrou S, Foelsch UR, Mathew CG, Lagerstrom-Fermer M, Schreiber S. Genetic variation in DLG5 is associated with inflammatory bowel disease. *Nat Genet* 2004;**36**:476–80.
21. **Daly MJ,** Pearce AV, Farwell L, Fisher SA, Latiano A, Prescott NJ, Forbes A, Mansfield J, Sand erson J, Langelier D, Cohen A, Bitton A, Wild G, Lewis CM, Annese V, Mathew CG, Rioux JD. Association of DLG5 R30Q variant with inflammatory bowel disease. *Eur J Hum Genet* 2005;**13**:835–9.
22. **Noble CL,** Nimmo ER, Drummond H, Smith L, Arnott ID, Satsangi J. DLG5 variants do not influence susceptibility to inflammatory bowel disease in the Scottish population. *Gut* 2005;**54**:1416–20.
23. **Török HP,** Glas J, Tonenchi L, Lohse P, Muller-Myhsok B, Limbersky O, Neugebauer C, Schnitzler F, Seiderer J, Tillack C, Brand S, Brunnler G, Jagiello P, Epplen JT, Griga T, Klein W, Schiemann U, Folwaczny M, Ochsenkuhn T, Folwaczny C. Polymorphisms

in the DLG5 and OCTN cation transporter genes in Crohn's disease. *Gut* 2005;**54**:1421–7.

24. **Vermeire S,** Pierik M, Hlavaty T, Claessens G, van Schuerbeeck N, Joossens S, Ferrante M, Henckaerts L, Bueno de Mesquita M, Vlietinck R, Rutgeerts P. Association of organic cation transporter risk haplotype with perianal penetrating crohn's disease but not with susceptibility to IBD. *Gastroenterology* 2005;**129**:1845–53.

25. **Büning C,** Geerdts L, Fiedler T, Gentz E, Pitre G, Reuter W, Luck W, Buhner S, Molnar T, Nagy F, Lonovics J, Dignass A, Land t O, Nickel R, Genschel J, Lochs H, Schmidt HH, Witt H. DLG5 variants in inflammatory bowel disease. *Am J Gastroenterol* 2006;**101**:786–92.

26. **Ferraris A,** Torres B, Knafelz D, Barabino A, Lionetti P, de Angelis GL, Iacono G, Papadatou B, D'Amato G, Di Ciommo V, Dallapiccola B, Castro M. Relationship between CARD15, SLC22A4/5 and DLG5 polymorphisms and early-onset inflammatory bowel diseases: an Italian multicentric study. *Inflamm Bowel Dis* 2006;**12**:355–61.

27. **Lakatos PL,** Fischer S, Claes K, Kovacs A, Molnar T, Altorjay I, Demeter P, Tulassay Z, Palatka K, Papp M, Rutgeerts P, Szalay F, Papp J, Vermeire S, Lakatos L. DLG5 R30Q is not associated with IBD in Hungarian IBD patients but predicts clinical response to steroids in Crohn's disease. *Inflamm Bowel Dis* 2006;**12**:362–8.

28. **Medici V,** Mascheretti S, Croucher PJ, Stoll M, Hampe J, Grebe J, Sturniolo GC, Solberg C, Jahnsen J, Moum B, Schreiber S, Vatn MH. Extreme heterogeneity in CARD15 and DLG5 Crohn disease-associated polymorphisms between German and Norwegian populations. *Eur J Hum Genet* 2006;**1**:459–68.

29. **Newman WG,** Gu X, Wintle RF, Liu X, van Oene M, Amos CI, Siminovitch KA. DLG5 variants contribute to Crohn disease risk in a Canadian population. *Hum Mutat* 2006;**27**:353–8.

30. **Tremelling M,** Waller S, Bredin F, Greenfield S, Parkes M. Genetic variants in TNF-alpha but not DLG5 are associated with inflammatory bowel disease in a large United Kingdom cohort. *Inflamm Bowel Dis* 2006;**12**:178–84.

31. **Biank V,** Friedrichs F, Babusukumar U, Wang T, Stoll M, Broeckel U, Kugathasan S. DLG5 R30Q variant is a female-specific protective factor in pediatric onset Crohn's disease. *Am J Gastroenterol* 2007;**102**:391–8.

32. **Cucchiara S,** Latiano A, Palmieri O, Staiano AM, D'Inca R, Guariso G, Vieni G, Rutigliano V, Borrelli O, Valvano MR, Annese V. Role of CARD15, DLG5 and OCTN genes polymorphisms in children with inflammatory bowel diseases. *World J Gastroenterol* 2007;**13**:1221–9.

33. **Pearce AV,** Fisher SA, Prescott NJ, Onnie CM, Pattni R, Green P, Forbes A, Mansfield J, Sand erson J, Schreiber S, Lewis CM, Mathew CG. Investigation of association of the DLG5 gene with phenotypes of inflammatory bowel disease in the British population. *Int J Colorectal Dis* 2007;**22**:419–24.

34. **Browning BL,** Huebner C, Petermann I, Demmers P, McCulloch A, Gearry RB, Barclay ML, Shelling AN, Ferguson LR. Association of DLG5 variants with

35. **Friedrichs F,** Brescianini S, Annese V, Latiano A, Berger K, Kugathasan S, Broeckel U, Nikolaus S, Daly MJ, Schreiber S, Rioux JD, Stoll M. Evidence of transmission ratio distortion of DLG5 R30Q variant in general and implication of an association with Crohn disease in men. *Hum Genet* 2006;**119**:305–11.

36. **Yamazaki K,** Takazoe M, Tanaka T, Ichimori T, Saito S, Iida A, Onouchi Y, Hata A, Nakamura Y. Association analysis of SLC22A4, SLC22A5 and DLG5 in Japanese patients with Crohn disease. *J Hum Genet* 2004;**49**:664–8.

37. **The International HapMap Consortium**. A haplotype map of the human genome. *Nature* 2005;**437**:1299–320.

38. **de Ridder L,** Weersma RK, Dijkstra G, van der Steege G, Benninga MA, Nolte IM, Taminiau JA, Hommes DW, Stokkers PC. Genetic susceptibility has a more important role in pediatric-onset Crohn's disease than in adult-onset Crohn's disease. *Inflamm Bowel Dis* 2007;**13**:1083–92.

39. **R Development Core Team**. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2006.

40. **Cochran WG.** The combination of estimates from different experiments. *Biometrics* 1954;**10**:101–129.

41. **Higgins JP,** Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;**21**:1539–58.

42. **DerSimonian R,** Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;**7**:177–88.

43. **Wigginton JE,** Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* 2005;**76**:887–93.

44. **Binder V.** Epidemiology of IBD during the twentieth century: an integrated view. *Best Pract Res Clin Gastroenterol* 2004;**18**:463–79.

45. **Gunderson KL,** Kuhn KM, Steemers FJ, Ng P, Murray SS, Shen R. Whole-genome genotyping of haplotype tag single nucleotide polymorphisms. *Pharmacogenomics* 2006;**7**:641–8.

46. **Corrao G,** Tragnone A, Caprilli R, Trallori G, Papi C, Andreoli A, Di Paolo M, Riegler G, Rigo GP, Ferrau O, Mansi C, Ingrosso M, Valpiani D. Risk of inflammatory bowel disease attributable to smoking, oral contraception and breastfeeding in Italy: a nationwide case-control study. Cooperative Investigators of the Italian Group for the Study of the Colon and the Rectum (GISC). *Int J Epidemiol* 1998;**27**:397–404.

47. **Garcia Rodriguez LA,** Gonzalez-Perez A, Johansson S, Walland er MA. Risk factors for inflammatory bowel disease in the general population. *Aliment Pharmacol Ther* 2005;**22**:309–15.

48. **Purmonen S,** Ahola TM, Pennanen P, Aksenov N, Zhuang YH, Tuohimaa P, Ylikomi T. HDLG5/KIAA0583, encoding a MAGUK-family protein, is a primary progesterone target gene in breast cancer cells. *Int J Cancer* 2002;**102**:1–6.

49. **Logan RF.** Inflammatory bowel disease incidence: up, down or unchanged? *Gut* 1998;**42**:309–11.