

ARTICLE

An association between the *PTGS2* rs5275 polymorphism and colorectal cancer risk in families with inherited non-syndromic predisposition

Jason Ross^{*,1,2}, Linda Lockett^{1,2}, Diana Brookes^{1,2}, Bruce Tabor^{1,3}, Konsta Duesing^{1,2}, Michael Buckley^{1,3}, Trevor Lockett^{1,2}, Peter Molloy^{1,2}, Finlay Macrae^{4,5}, Graeme Young⁶, Ignacio Blanco⁷, Gabriel Capella⁷ and Garry N Hannan^{1,2}

Recently our group completed a genome-wide linkage study investigating Australian and Spanish families with inherited risk of colorectal cancer (CRC). A minor linkage peak from that study located on chromosome 1 correlates with the location of a known CRC risk-modifying gene, prostaglandin synthase (*PTGS2*). *PTGS2* encodes the inducible prostaglandin synthase enzyme cyclooxygenase-2 (COX-2). Prostaglandins are implicated in the initiation of carcinogenesis and progression of tumours. Sequencing of *PTGS2* in a small subset of affected individuals identified a high frequency of the minor C allele of single nucleotide polymorphism rs5275. We then genotyped the rs5275 polymorphism in 183 affected and 223 unaffected individuals from our CRC predisposed families. Tests for association in the presence of linkage were made using family-based association tests. The C allele was found to be significantly associated ($P < 0.01$) with diagnosis of hereditary non-syndromic CRC ($P = 0.0094$, dominant model) and an earlier age of diagnosis ($P = 0.0089$, heterozygous-advantage model). Interestingly, by stratifying the age of diagnosis data, we observed a speculative gender-discordant effect. Relative to other groups, female CC carriers were diagnosed less when young, but by 60 years of age were the most at risk group. Conversely, CT carriers of both genders showed a consistently earlier diagnosis relative to TT carriers. Our results suggest potential differential age- and gender-dependent efficacies of chemopreventative COX-2 inhibitors in the context of non-syndromic colorectal cancer.

European Journal of Human Genetics (2013) 21, 1389–1395; doi:10.1038/ejhg.2013.53; published online 27 March 2013

Keywords: *PTGS2*; COX-2; rs5275; colorectal; cancer; hereditary

INTRODUCTION

Prostaglandin-endoperoxide G/H synthase 2, also known as cyclooxygenase-2 (COX-2), is encoded by the gene *PTGS2*. COX-2 functions as a homodimer and converts arachidonic acid to prostaglandin H₂. In turn, this is converted to other pro-inflammatory prostaglandins such as prostaglandin E₂ (PGE₂). There is a causal relationship between chronic inflammation and cancer,¹ with PGE₂ known to stimulate cancer progression.² In particular, inflammatory diseases of the colon such as Crohn's disease and ulcerative colitis are associated with increased colorectal cancer (CRC) risk.³ Aspirin and other non-steroidal anti-inflammatory inhibitors of COX-2 are known to act as CRC chemopreventatives.⁴ COX-2 is known to be overexpressed in ~85% of human colorectal carcinomas, 50% of colorectal adenomas,⁵ as well as other epithelial and nonepithelial tumours (reviewed in Koki and Masferrer⁶). There is evidence that COX-2 overexpression is also involved in the aetiology of familial adenomatous polyposis (FAP), a CRC predisposition syndrome caused by an inherited mutation of the Adenomatous polyposis coli (*APC*) gene. The polyps removed from FAP patients show overexpression of COX-2 in the fibroblast and endothelial cells.⁷ In a FAP animal model, COX-2 overexpression has a direct role in the formation of tumours as polyposis in *Apc*-knockout

mice is suppressed by the inhibition of COX-2.⁸ It has also been reported that FAP patients experience significantly less adenoma recurrence when taking the COX-2 inhibitor drug, celecoxib.⁹ High COX-2 expression is also observed in another familial CRC syndrome, MUTYH-associated polyposis.¹⁰

As COX-2 overexpression can be oncogenic, genetic variants that alter the expression of COX-2 may influence development of CRC. A large number of single nucleotide polymorphisms (SNPs) in and around the *PTGS2* locus have been reported. If these genetic variations alter the production or function of the COX-2 enzyme, they may modulate the risk of cancer. Non-synonymous coding SNPs within the exons of *PTGS2* are very rare. However, the relatively common non-coding SNPs rs20417 (−765G>C) in the promoter and rs5275 (+8473T>C) in the 3'-UTR of the gene have been shown to reduce promoter activity¹¹, and alter mRNA stability,^{12,13} respectively. The rs5275 SNP is located within the microRNA miR-542-3p binding site of the COX-2 transcript, with the C allele interfering with microRNA binding. Transcripts bearing the C allele exhibit less microRNA-mediated transcript degradation, increased mRNA stabilisation and thereby direct more COX-2 production.¹³

¹CSIRO, Preventative Health National Research Flagship, North Ryde, NSW, Australia; ²Animal, Food and Health Sciences, CSIRO, North Ryde, NSW, Australia; ³Mathematics, Informatics and Statistics, CSIRO, North Ryde, NSW, Australia; ⁴Department of Medicine, University of Melbourne, Parkville, VIC, Australia; ⁵Colorectal Medicine and Genetics, The Royal Melbourne Hospital, Parkville, VIC, Australia; ⁶Flinders Centre for Innovation in Cancer, Flinders University, Bedford Park, SA, Australia; ⁷Hereditary Cancer Program, Genetic Counselling Unit, Institut Català d'Oncologia-IDIBELL and University of Barcelona, Barcelona, Spain

*Correspondence: Dr J Ross, CSIRO Animal, Food and Health Sciences, CSIRO, 11 Julius Avenue, P.O. Box 52, North Ryde, NSW 1670, Australia. Tel: +61 294 905 015; Fax: +61 294 905 020; E-mail: jason.ross@csiro.au

Received 15 November 2012; revised 15 February 2013; accepted 19 February 2013; published online 27 March 2013

The rs5275 C allele has a frequency of ~ 0.37 in Caucasian populations, with a higher frequency in African (~ 0.67) and less in Asian (~ 0.17) populations (dbSNP, submission ss66858504). This SNP has been well studied in many disease contexts and cancer in particular. There are reported significant associations with basal cell, bladder, breast, colorectal, gastric, lung, prostate and ovarian cancers.^{14–28} Across cancer tissues of different origin, each allele paradoxically has been associated with increased risk of cancer. However, if studies are stratified by cancer type, there is greater consistency in the assignment of the risk allele.²⁹ Other biological factors such as gender and age and environmental factors such as non-steroidal antiinflammatory drug use and smoking status also influence the strength of association with cancer incidence.^{15,25,27,28,30,31} Considered together, the prevailing evidence suggests that rs5275 is associated with cancer incidence, but the associations are complex.

In our previous linkage study of families with inherited non-syndromic CRC predisposition, we identified weak linkage to the chromosome 1 region containing the *PTGS2* gene locus.³² In this present study we chose to sequence the coding region of *PTGS2* in a subset of 16 CRC-affected people with linkage to the chromosome 1 region. The sequencing suggested novel or rare coding SNPs of rs5275 were not contributing to disease. However, we noted a relatively high frequency of the C allele of the common SNP, rs5275. We then genotyped rs5275 in all family members in the Australian and Spanish families previously collected for the linkage study, and looked for an association with CRC diagnosis and age of diagnosis.

MATERIALS AND METHODS

Study population

This family-based association study examined non-syndromic, high-risk CRC families. Families with known colon cancer syndromes, including HNPCC or Lynch syndrome, FAP, hereditary mixed polyposis, juvenile polyposis, Peutz–Jeghers, Cowden's syndrome or *MUTYH*-associated polyposis were excluded. A description of these families and criteria for inclusion in the study have been described previously.³² Briefly, these families are defined as those containing at least one affected person who has one or more first-degree affected relative(s) and where the known causal mutations had been excluded. As in our previous study, affected status was defined as diagnosis with either colorectal carcinoma or advanced adenoma. In turn, advanced adenoma was defined as three or more synchronous or metachronous adenomas and/or adenoma(s) with villous morphology, and/or with severe dysplasia, and/or diameter ≥ 10 mm. The cohort is mostly sibships with missing parental genotypes and is described further in Table 1. In some instances extended pedigrees were genotyped; these include families of up to four generations (including non-genotyped founders). A number of the pedigrees contained parent–child, avuncular or cousin affected pairs. The families show weak genetic linkage with CRC (non-parametric LOD = 1.71) in a 21 Mb region contained within chromosome 1q25.2–q32.1 with the *PTGS2* gene locus situated in the middle of this region. The population examined comprises 418 genotyped individuals, 241 described in the linkage study and a further 177 individuals from cohort families, which were considered uninformative for our earlier non-parametric linkage analysis. In total there were 183 affected cases (179 had a recorded age of onset) and 223 unaffected relatives, mostly sibs. An additional 12 people without CRC diagnosis, but under 55 years of age or with previous detection of colorectal polyps were considered to have unknown disease status. All people in the study had their age at blood draw recorded. There were 173 males and 245 females in the study (Table 1).

Variant screening

Genomic DNA was extracted from white blood cells using a PAXgene Blood DNA Kit (Qiagen, Melbourne, Australia). DNAs were amplified, labelled using BigDye Terminator v3.1 chemistry and run on a 3730xl capillary sequencer

Table 1 Family statistics

Family statistics		
<i>Families (including founders)</i>		
Discrete pedigrees		147
Average family size		4.34
Average generations		1.91
<i>Origin</i>		
Flinders centre for cancer prevention and control		60
Royal Melbourne hospital		45
Institut Català d'Oncologia		11
<i>Gender</i>		
Males		173
Females		245
<i>Median age</i>		
Affected		55 (23–85)
Unaffected		57 (21–93)
<i>Genotyped</i>		
Total		418
Affection status known		406
<i>Disease status</i>		
Affected		183
Unaffected		223
Unknown		12
<i>Pairs</i>	<i>Both affected</i>	<i>Discordant</i>
Sib pairs	65	225
Cousin pairs	1	6
Parent–child pairs	23	61
Grandparent–child pairs	0	2
Avuncular pairs	12	48

(Applied Biosystems, Foster City, CA, USA). Primers were selected using the VariantSEQR (Applied Biosystems) resequencing set for *PTGS2* (with sequences provided in Supplementary Table 1). Sequences were inspected visually for variants.

Genotyping

To ensure exceptionally high genotyping quality, the rs5275 (8473T>C) genotype was determined for every possible individual using both the PCR-based primer-introduced restriction analysis (PIRA) method as described by Hu *et al*²⁴ and the TaqMan 5' nuclease allelic discrimination assay (Applied Biosystems). Two individuals could only be genotyped using the PIRA assay owing to small quantities of available DNA.

TaqMan assays were performed in 25 μ l volumes with 10 ng of template DNA. TaqMan probe sequences and thermal cycling conditions were as described in the variantGPS database (<http://variantgps.nci.nih.gov>) under polymorphism ID, 'CGFID_4979'; otherwise, conditions were as described by the manufacturer. The genotyping assays were performed in 96-well plates in a LightCycler 480 real-time PCR system (Roche Applied Sciences, Sydney, Australia). For PIRA assays, 50 ng of template DNA was amplified (in duplicate) and resultant PCR products digested overnight using BclI (New England Biolabs Inc., Beverly, MA, USA). The digested PCR products were separated on a Criterion TBE 10% polyacrylamide gels (Bio-Rad Laboratories, Sydney, Australia) and resolved using Gel Red Nucleic Acid Stain (Biotium, Hayward, CA, USA) and a Typhoon 9400 scanner (GE Healthcare, Sydney, Australia).

Statistical analysis

A survival model was fitted and Kaplan–Meier plots drawn in R (version 2.14, 64-bit) using the ‘survival’ package. Generalised family-based association tests (FBAT)³³ implemented in the PBAT version 3.6 software³⁴ were carried out using generalised estimating equations and *P*-values generated from the asymptotic Normal distribution. Genotypes were analysed using both dichotomous affection status and censored time-to-onset analyses. For the time-to-onset analysis, we used a Wilcoxon Logrank FBAT statistic. The test used a null hypothesis of linkage and no association with sandwich variance estimation. Pedigree correlation estimates can become unstable in large pedigrees. Sandwich variance is useful for robustly estimating correlation between members in these large pedigrees. These analyses were also repeated with consideration of gender. FBAT analyses using nuclear families were implemented in the FBAT V2.0.4 beta software (<http://www.hsph.harvard.edu/fbat/fbat.htm>).³⁵ The generalised disequilibrium test V0.1.1 software³⁶ was used to test for association in dichotomous relative pairs.

RESULTS

Variant screening

DNAs from 16 people diagnosed with CRC and good evidence for linkage to the chromosome 1 region were forward and reverse capillary sequenced across the exons, 3-UTR and promoter region of *PTGS2*. In total, 16 SNPs were discovered, of which two were novel (Supplementary Table 2). For each novel SNP, there was only a single

observation, so it is highly unlikely that they are the cause of inherited cancer predisposition in these 16 people. Given this, we did not follow up these observations in a larger cohort. Interestingly, we observed the SNP rs5275 C variant allele was enriched in frequency compared with Caucasian population estimates (~0.394; dbSNP135), with 77% of people in the sample having the CT or CC genotype, giving a C allele frequency of 0.462. Given this increased C allele frequency and previous reports of an association of rs5275 with CRC diagnosis, we chose to further analyse this SNP in the whole cohort. A single SNP should be a reasonable marker for the involvement of *PTGS2* with CRC as the gene is particularly small, only occupying 8.6 kb of space on chromosome 1. This small locus size implies other genetic variations within *PTGS2* should be in high linkage disequilibrium.

Genotyping and quality control

Across 418 individuals, all except two were genotyped for rs5275 using both the PIRA and TaqMan methods. There were 20 discordant genotypes between the two technologies. In our hands, Taqman was clearly the more accurate genotyping method—all 20 discordant genotypes were resolved as PIRA genotyping errors (following repeat *BclI* restriction digest and genotyping on another polyacrylamide gel). We found the C allele frequency to be 0.334, which is consistent with other Caucasian populations. The rs5275 genotype was found to be in Hardy-Weinberg equilibrium.

Table 2 Results of rs5275 genotyping split by disease status and gender

Affection status	Gender	Genotype			Total
		CC	CT	TT	
Affected	Male	4 (47.5)	42 (56) ^a	34 (58)	80 (56)
	Female	12 (60) ^a	51 (53.5) ^a	40 (52.5)	103 (55)
Unaffected	Male	16 (55)	31 (58)	42 (57.5)	89 (57)
	Female	17 (49)	52 (61)	65 (58)	134 (58)
Unknown	Male	0	1	3	4
	Female	0	4	4	8
Total		49 (54)	181 (57)	188 (57)	418 (57)

Median ages are given in brackets. For affected people this is the median age of disease diagnosis and for unaffected people the age at blood draw. Four affected people did not contribute to the median age of disease diagnosis as the data were not available. These included one affected female CC carrier and two male and one female affected CT carriers.
^aSome missing age of colorectal cancer diagnosis data.

Summary statistics and survival analysis

Summary statistics subdivided by disease status and gender are presented in Table 2. The median age of disease diagnosis across the genotypes was 55.5, 55.0 and 58.0 for the TT, CT and CC genotypes, respectively. Interestingly, there is some discordance between males and females in the median age of disease diagnosis (Table 2). There is evidence that males carrying at least one C allele have an earlier age of disease diagnosis, while the opposite is observed for females. This is illustrated in more detail in Figures 1a and b. In Kaplan–Meier curves (Figures 2a and b) carriers of the heterozygous CT genotype in both genders trended toward an earlier age of diagnosis than the more common TT genotype carriers. For the 12 females diagnosed with CRC carrying the rare CC genotype, the trend suggests the risk imposed by the CC genotype may vary with age. Until ~55 years of age female CC carriers have the lowest rate of cancer diagnosis; however, within the following 5 years the proportion of carriers diagnosed with CRC is comparable with that of male CT

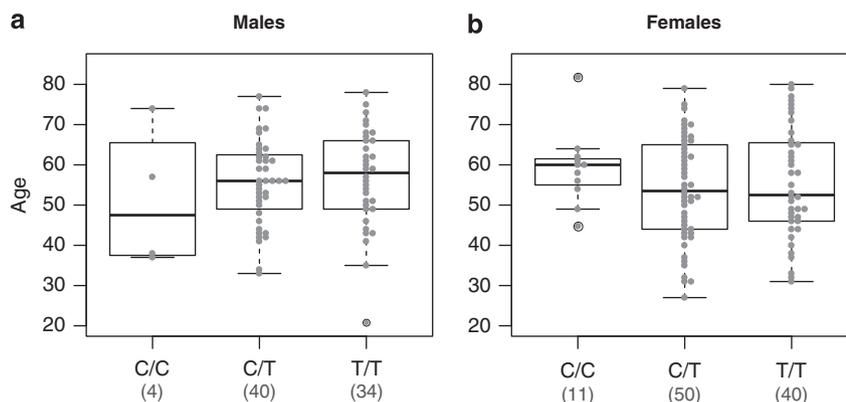


Figure 1 Combination boxplot-stripchart by genotype and gender. (a) Affected male C allele carriers have an earlier median age of CRC diagnosis, while the opposite is observed for (b) female C allele carriers. The box represents the upper and lower quartiles and the intersecting horizontal line, the median. The filled grey points are individual patient ages of diagnosis with multiple cases stacked horizontally. Below the genotype the number in parentheses is the total number of cases observed for that genotype with a reported age of CRC diagnosis.

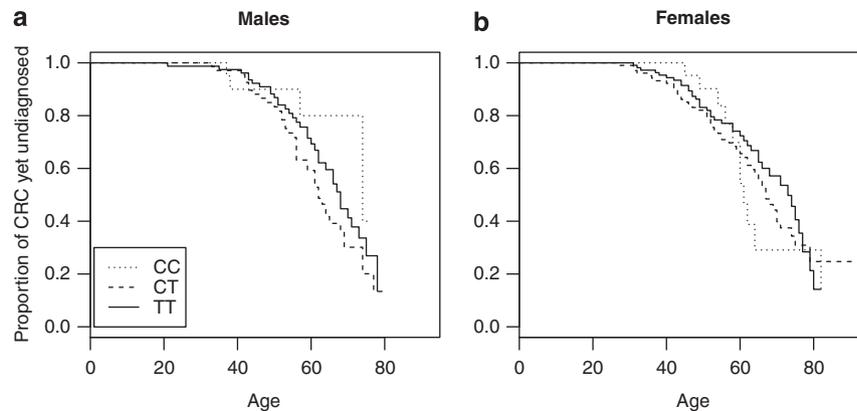


Figure 2 Kaplan–Meier analysis of ages of diagnosis, with consideration of rs5275 genotype and gender. The survival function of right-censored age of diagnosis data for (a) 78 males and (b) 101 females were formed with a Kaplan–Meier estimator and genotype and gender as factors. The estimates are plotted by genotype as a function of the proportion of CRC as yet undiagnosed at a given age. As illustrated in the legend on panel (a), the dotted, dashed and solid lines represent the CC, CT and TT genotype estimates, respectively. Crosses on the estimator line are uncensored people—so the age at blood draw of relatives without evidence of CRC.

genotype carriers. Little can be stated regarding male CC carriers as there were only four observations. The apparent trends in the data were tested using a Cox proportional hazard model conditioned on the pedigrees and stratified by gender. Without clustering there was evidence of a violation of the proportional hazards assumption within pedigrees ($P = 0.026$), with the effect of pedigree increasing over time. To account for the non-independent observations within families, we clustered the data within pedigrees. The odds ratios (OR), with CC as the reference, for the TT and CT genotypes were 1.02 and 1.32, respectively (Likelihood ratio test, $P = 0.221$; Wald test, $P = 0.168$). Although the result is non-significant, the increased OR for the CT genotype, but not the CC or TT genotypes was interesting. We followed this up by collapsing the genotype into heterozygotes (CT genotype) and homozygotes (CC or TT genotypes). In this instance, homozygotes were found to be protected against CRC (OR = 0.78) with the difference in risk approaching significance (Likelihood ratio test, $P = 0.101$; Wald test, $P = 0.069$). Although the survival analysis was not significant, it suggests tests of association with CRC age of diagnosis should have the most power under a heterozygous-advantage genetic model.

Family-based association testing

FBATs allow determination of associations in the presence of intra-family correlations that result from allelic bias and the tests are robust to population stratification. In particular, the classes of FBATs implemented in the PBAT software are well suited for this current study, as the tests can generalise to large or small arbitrary pedigrees and to pedigrees with missing parental genotypes.³⁴ Further, PBAT can handle both dichotomous and censored age of diagnosis phenotypes. The PBAT FBAT showed the C allele of rs5275 as significantly associated with the diagnosis of CRC in these predisposed families (Table 3). The additive (increasing risk with increasing C dosage, $P = 0.0159$), dominant (CT and CC impose increased risk, $P = 0.0094$) and heterozygous-advantage (only CT imposes risk, $P = 0.0496$) models all show significant association between carriers of the minor C allele and the diagnosis of CRC. We also considered age at CRC diagnosis as a phenotype and found similar results (Table 4). The minor C allele was significantly associated with an earlier age of diagnosis under the heterozygous-advantage genetic model ($P = 0.0089$), the dominant model

Table 3 Results of the family-based association testing between diagnosis of colorectal cancer and rs5275 genotype

Model	Informative families		PBAT (diagnosis)	
	T allele	C allele	T allele	C allele
Additive	35	35	0.0159^a	0.0159
Dominant	23	30	0.9255 ^a	0.0094
Recessive	24	10	0.0055^a	0.9255
Heterozygous-advantage	35	35	0.0496	0.0496

Results are grouped by genetic model under consideration. P -values for each allele were generated by PBAT using a generalised estimating equation for dichotomous traits. Significant ($P < 0.05$) results are in bold and results with evidence of a negative association are denoted with an asterisk after the P -value. For each model, the number of informative families represents the number of families without homozygous parents (or evidence to support this based upon the sufficient statistic).

^aA negative correlation was observed between the phenotype and the number of transmitted alleles.

($P = 0.0116$) and additive model ($P = 0.0486$) using an FBAT–Wilcoxon statistic. Considering the gender-discordant trends highlighted in the survival analysis, we also stratified the age of diagnosis association analysis by gender (Table 4). Although there were considerably less subjects in each test, in the more populous female group the association under the dominant ($P = 0.0451$) and heterozygous-advantage ($P = 0.0427$) models were still significant. While no significant association was observed in males, in a similar fashion, the dominant ($P = 0.1886$) and heterozygous-advantage ($P = 0.1605$) models were also the best supported.

DISCUSSION

Our previous linkage study suggested linkage between familial inherited non-syndromic CRC and a region of chromosome 1 spanning the *PTGS2* gene locus.³² There is extensive evidence to implicate the involvement of the *PTGS2* locus gene product, COX-2, in carcinogenesis (Mann and DuBois³⁷) and many genetic association and biochemical studies support the association of rs5275 with CRC diagnosis or aetiology. We confirm and extend these earlier studies, reporting here the first study to consider the association between rs5275 and cancer diagnosis within predisposed families. We show good evidence that rs5275 genotype is significantly associated with

Table 4 Results of the family-based association testing between age of diagnosis of colorectal cancer and rs5275 genotype

Model	Informative families		PBAT (age of diagnosis)	
	T allele	C allele	T allele	C allele
<i>All people</i>				
Additive	35	35	0.0486^a	0.0486
Dominant	23	30	0.3334	0.0116
Recessive	24	10	0.0109^a	0.3334 ^a
Heterozygous-advantage	35	35	0.0089	0.0089
<i>Males only</i>				
Additive	30	30	0.3744 ^a	0.3744
Dominant	20	27	0.6168	0.1886
Recessive	21	9	0.1663 ^a	0.6168 ^a
Heterozygous-advantage	30	30	0.1605	0.1605
<i>Females only</i>				
Additive	33	33	0.0912 ^a	0.0912
Dominant	23	28	0.4340	0.0451
Recessive	22	10	0.0508 ^a	0.4340 ^a
Heterozygous-advantage	33	33	0.0427	0.0427

Three sets of results are presented, the FBAT–Wilcoxon statistic *P*-values for all people followed by stratification into males and females only. In each instance the results are grouped row-wise by genetic model and column-wise by allele. Like Table 3, significant ($P < 0.05$) results are in bold, negative associations denoted with an asterisk and informative families defined in the same manner.

^aA negative correlation was observed between the phenotype and the number of transmitted alleles.

both diagnosis and age of diagnosis of non-syndromic CRC, with carriers of the rs5275 variant C allele having a higher risk of developing CRC ($P < 0.01$). This association is consistent with previous reports considering sporadic CRC, which also report a significant association of the C allele with increased cancer risk.^{18,19,21}

We found consideration of the age of diagnosis to generate some interesting findings. While age is a major risk factor for all CRC, Kaplan–Meier plots suggest a possible gender-discordant effect with female carriers of the rare CC genotype having a reduced rate of CRC diagnosis when young, but increased diagnosis of familial CRC in later years. We state this tentatively, as there were only 12 observations to support the finding. Although the trend observed in age of diagnosis for CC carriers is complex, carriers of the CT genotype in both genders seem to consistently have an earlier age of onset than carriers of the TT genotype. To test for differences in age of diagnosis between genotypes, we used a likelihood ratio test with clustering by pedigree. Although this test was not significant, the simple survival analysis we used has caveats; it does not account for degrees of relatedness between members of a pedigree, nor does it consider the genetic model or correlation due to genetic linkage. To account for this, we used PBAT that can model as an association in the presence of genetic linkage with consideration of the genetic model and full pedigree structure. In this instance, we found a highly significant association ($P < 0.01$) between the C allele and both diagnosis and age of diagnosis of CRC. Although the dominant genetic model was well supported for both the diagnosis and age of diagnosis phenotype ($P = 0.0094$ and $P = 0.0116$, respectively), we note the heterozygous-advantage model was less well supported for the diagnosis phenotype ($P = 0.0496$, compared with $P = 0.0089$), while the additive model was less well supported for the age of diagnosis phenotype ($P = 0.0486$, compared with $P = 0.0159$). This discrepancy can be

explained by the difference in statistic used to test the two phenotypes and by consideration of the Kaplan–Meier plot. The FBAT–Wilcoxon statistic used in the age of diagnosis analysis assigns more weight to early CRC diagnoses, while the FBAT statistic for dichotomous traits considers all diagnoses equally with a simple binary encoding.³⁸ In young females, CT carriers have the highest rate of CRC diagnosis so there is an expectation of good support for a heterozygous-advantage model using an FBAT–Wilcoxon statistic.

Across the two phenotypes, the consistency we observe in the support for an association with the C allele under the additive, dominant and heterozygous-advantage genetic models is encouraging. The age of diagnosis is a more information-rich phenotype than diagnosis alone; however, specification of the age of diagnosis is prone to some vagaries and bias. Foremost, the progression of CRC from early- to late-stage disease can take multiple years, and patients are diagnosed at various stages of disease progression. Additionally, in a familial cancer setting, often a diagnosis in the proband may lead to earlier surveillance of the proband's siblings and offspring. Finally, there is the potential for error in self-reporting an age of diagnosis. Although age of diagnosis is a more information-rich phenotype, a simple dichotomous trait will avoid these caveats, as it is a test of 'if' and not 'when' a person will get disease.

Consideration of the association testing and survival analysis together suggests that the association of rs5275 genotype with non-syndromic CRC is complex and has a possible gender effect. The tentative finding that female carriers of the CC genotype are protected from CRC development early in their lives, but show increased risk around 55–60 years of age, is particularly interesting. Considering the median age of menopause is estimated to be 51.3 years,³⁹ this increased risk may correlate with prior onset of menopause. Age discordant and menopause-dependent effects in women's predisposition towards CRC diagnosis and prognosis is well documented.⁴⁰ It is also known that hormone replacement therapy is protective against the onset of CRC in women, which suggests that reduced oestrogen production postmenopause may increase the risk of CRC.⁴¹ Functional evidence exists for the involvement of oestrogen in CRC. Oestrogen has access to normal colon epithelium, with oestrogen receptor beta (ER β) abundantly expressed in normal colonic epithelia in both genders; however, in CRC cells this expression of ER β is significantly reduced.^{42,43} We postulate that the observed age-dependent risk imposed in females by the CC genotype is due to the protective role of oestrogen in younger women.

Our findings of a complex association are supported by comparisons to the published case-control studies showing significant association of rs5275 with cancer risk (further discussion is available in the Supplementary Material). Collectively, the differences in findings in these studies suggest the allele associated with risk is dependent upon factors, such as the cancer tissue of origin, environmental influences and ethnicity. If only CRC association studies showing significant association are considered, a more consistent trend emerges. A study by Ali *et al*¹⁹ with 726 Caucasian cases and 729 controls, found a complex gender discordant association similar to this present study. In males, carriers of the CT genotype, but not CC genotype, were at increased risk of CRC with a respective OR of 1.31 and 1.02. In females, the trend was reversed (with OR = 0.94 and OR = 1.40). The two other sporadic CRC case-control studies reporting significant associations with rs5275 genotype are from the same group and do not explicitly segregate the data by gender.^{18,21} However, if we consider the associations found within this pair of Caucasian cohorts and relate this to the proportion of males in each,

there is some evidence of a gender-discordant effect. In the first cohort with 53.3% male cases, only the CT genotype was found to be associated with CRC incidence (OR = 1.47).²¹ In the second cohort with 22.4% male cases, only the CC genotype was significantly associated with CRC (OR = 1.14). We suggest that the gender-discordant risk of carrying the C allele is responsible for the observation of an apparent heterozygous-advantage model in the mostly male cohort.

Given the complexity of the association, we also employed other association tests to test the sensitivity of the result (Supplementary Material). Only the PBAT algorithm that considers affected people across a number of genetic models and makes use of full pedigree structures was sensitive enough to find an association. Our results also suggest that further family-based studies of rs5275 would benefit from an increased number of families in the analysis.

Given that the C allele leads to increased mRNA stabilisation of the *PTGS2* transcript,⁴⁴ the association we observe between the rs5275 C allele and non-syndromic CRC suggests that tumours of C allele carrying patients may have overexpression of COX-2, similar to patients with FAP and MAP syndromes. It would be interesting to measure COX-2 expression levels in colorectal tissue and tumours from patients with inherited non-syndromic CRC predisposition, and to correlate this against their rs5275 genotype.

This current study strengthens the evidence for a link between rs5275 genotype and risk of diagnosis of CRC. We also establish an association between rs5275 genotype and age of diagnosis of CRC. This is the first study of rs5275 in a familial cancer setting, and the evidence for association suggests that rs5275 genotype should be further investigated in this context in larger studies. Our use of gender as a factor and age of diagnosis as a phenotype reveal complexities in the association with rs5275 genotype, this in turn suggests that chemopreventative COX-2 inhibitor drugs may have different efficacies depending upon age, gender and rs5275 genotype. We suggest consideration of gender may help explain the contrary results in the literature for the protection associated with the 'C' allele in the breast¹⁷ and ovarian cancer²⁸ and risk in prostate cancer.²⁶ Oestrogen activation of COX-2 is a compelling explanation for the tentative gender discordant and female age-dependent associations observed in this study. The evidence, in this and other studies, for the imposition of risk in carrying the rs5275 C allele being dependent upon gender and tissue suggests there would be a fairly low-selection pressure, and may explain why the C allele is relatively common in some populations. The inclusion of gender as a factor in future association studies, particularly of *PTGS2*, would be prudent.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We are grateful to the participating families in our linkage study. We also thank Dr Nicholas Archer and Dr Yalchin Oytam for critically reviewing this manuscript. This work is supported within CSIRO by the CSIRO Preventative Health National Research Flagship and at the ICO by contract/grant sponsor: Asociación Española Contra el Cáncer and PI10/00748.

- 4 Cuzick J, Otto F, Baron JA *et al*: Aspirin and non-steroidal anti-inflammatory drugs for cancer prevention: an international consensus statement. *Lancet Oncol* 2009; **10**: 501–507.
- 5 Eberhart CE, Coffey RJ, Radhika A, Giardiello FM, Ferrenbach S, DuBois RN: Up-regulation of cyclooxygenase 2 gene expression in human colorectal adenomas and adenocarcinomas. *Gastroenterology* 1994; **107**: 1183–1188.
- 6 Koki AT, Masferrer JL: Celecoxib: a specific COX-2 inhibitor with anticancer properties. *Cancer Control* 2002; **9**: 28–35.
- 7 Sonoshita M, Takaku K, Oshima M, Sugihara K, Taketo MM: Cyclooxygenase-2 expression in fibroblasts and endothelial cells of intestinal polyps. *Cancer Res* 2002; **62**: 6846–6849.
- 8 Oshima M, Dinchuk JE, Kargman SL *et al*: Suppression of intestinal polyposis in *Apc* delta716 knockout mice by inhibition of cyclooxygenase 2 (COX-2). *Cell* 1996; **87**: 803–809.
- 9 Bertagnoli MM, Eagle CJ, Zauber AG *et al*: Celecoxib for the prevention of sporadic colorectal adenomas. *N Engl J Med* 2006; **355**: 873–884.
- 10 Fratini M, Carnevali I, Signoroni S *et al*: Cyclooxygenase-2 expression in FAP patients carrying germ line MYH mutations. *Cancer Epidemiol Biomarkers Prev* 2005; **14**: 2049–2052.
- 11 Papafili A, Hill MR, Brull DJ *et al*: Common promoter variant in cyclooxygenase-2 represses gene expression: evidence of role in acute-phase inflammatory response. *Arterioscler Thromb Vasc Biol* 2002; **22**: 1631–1636.
- 12 Cok SJ, Morrison AR: The 3'-untranslated region of murine cyclooxygenase-2 contains multiple regulatory elements that alter message stability and translational efficiency. *J Biol Chem* 2001; **276**: 23179–23185.
- 13 Moore AE, Young LE, Dixon DA: A common single-nucleotide polymorphism in cyclooxygenase-2 disrupts microRNA-mediated regulation. *Oncogene* 2012; **31**: 1592–1598.
- 14 Vogel U, Christensen J, Wallin H, Friis S, Nexø BA, Tjønneland A: Polymorphisms in COX-2, NSAID use and risk of basal cell carcinoma in a prospective study of Danes. *Mutat Res* 2007; **617**: 138–146.
- 15 Yang H, Gu J, Lin X *et al*: Profiling of genetic variations in inflammation pathway genes in relation to bladder cancer predisposition. *Clin Cancer Res* 2008; **14**: 2236–2244.
- 16 Jung JH, Chae YS, Moon JH *et al*: TNF superfamily gene polymorphism as prognostic factor in early breast cancer. *J Cancer Res Clin Oncol* 2010; **136**: 685–694.
- 17 Cox DG, Buring J, Hankinson SE, Hunter DJ: A polymorphism in the 3' untranslated region of the gene encoding prostaglandin endoperoxide synthase 2 is not associated with an increase in breast cancer risk: a nested case-control study. *Breast Cancer Res* 2007; **9**: R3.
- 18 Siezen CL, Bueno-de-Mesquita HB, Peeters PH, Kram NR, van Doeselara M, van Kranen HJ: Polymorphisms in the genes involved in the arachidonic acid-pathway, fish consumption and the risk of colorectal cancer. *Int J Cancer* 2006; **119**: 297–303.
- 19 Ali IU, Luke BT, Dean M, Greenwald P: Allelic variants in regulatory regions of cyclooxygenase-2: association with advanced colorectal adenoma. *Br J Cancer* 2005; **93**: 953–959.
- 20 Kim JG, Chae YS, Sohn SK *et al*: Prostaglandin synthase 2/cyclooxygenase 2 (PTGS2/COX2) 8473T>C polymorphism associated with prognosis for patients with colorectal cancer treated with capecitabine and oxaliplatin. *Cancer Chemother Pharmacol* 2009; **64**: 953–960.
- 21 Siezen CL, van Leeuwen AI, Kram NR, Luken ME, van Kranen HJ, Kampman E: Colorectal adenoma risk is modified by the interplay between polymorphisms in arachidonic acid pathway genes and fish consumption. *Carcinogenesis* 2005; **26**: 449–457.
- 22 Canzian F, Franceschi S, Plummer M *et al*: Genetic polymorphisms in mediators of inflammation and gastric precancerous lesions. *Eur J Cancer Prev* 2008; **17**: 178–183.
- 23 Campa D, Zienolddiny S, Maggini V, Skaug V, Haugen A, Canzian F: Association of a common polymorphism in the cyclooxygenase 2 gene with risk of non-small cell lung cancer. *Carcinogenesis* 2004; **25**: 229–235.
- 24 Hu Z, Miao X, Ma H *et al*: A common polymorphism in the 3'UTR of cyclooxygenase 2/prostaglandin synthase 2 gene and risk of lung cancer in a Chinese population. *Lung Cancer* 2005; **48**: 11–17.
- 25 Park JM, Choi JE, Chae MH *et al*: Relationship between cyclooxygenase 8473T>C polymorphism and the risk of lung cancer: a case-control study. *BMC Cancer* 2006; **6**: 70.
- 26 Danforth KN, Hayes RB, Rodriguez C *et al*: Polymorphic variants in PTGS2 and prostate cancer risk: results from two large nested case-control studies. *Carcinogenesis* 2008; **29**: 568–572.
- 27 Pu X, Lippman SM, Yang H, Lee JJ, Wu X: Cyclooxygenase-2 gene polymorphisms reduce the risk of oral premalignant lesions. *Cancer* 2009; **115**: 1498–1506.
- 28 Lurie G, Terry KL, Wilkens LR *et al*: Pooled analysis of the association of PTGS2 rs5275 polymorphism and NSAID use with invasive ovarian carcinoma risk. *Cancer Causes Control* 2010; **21**: 1731–1741.
- 29 Zhu W, Wei BB, Shan X, Liu P: -765G>C and 8473T>C polymorphisms of COX-2 and cancer risk: a meta-analysis based on 33 case-control studies. *Mol Biol Rep* 2010; **37**: 277–288.
- 30 Vogel U, Christensen J, Wallin H *et al*: Polymorphisms in genes involved in the inflammatory response and interaction with NSAID use or smoking in relation to lung cancer risk in a prospective study. *Mutat Res* 2008; **639**: 89–100.
- 31 Siezen CL, Tijhuis MJ, Kram NR *et al*: Protective effect of nonsteroidal anti-inflammatory drugs on colorectal adenomas is modified by a polymorphism in peroxisome proliferator-activated receptor delta. *Pharmacogenomics* 2006; **16**: 43–50.

1 Coussens LM, Werb Z: Inflammation and cancer. *Nature* 2002; **420**: 860–867.

2 Castellone MD, Teramoto H, Williams BO, Druey KM, Gutkind JS: Prostaglandin E2 promotes colon cancer cell growth through a Gs-axin-beta-catenin signaling axis. *Science* 2005; **310**: 1504–1510.

3 Bernstein CN, Blanchard JF, Kliever E, Wajda A: Cancer risk in patients with inflammatory bowel disease: a population-based study. *Cancer* 2001; **91**: 854–862.

- 32 Saunders IW, Ross J, Macrae F *et al*: Evidence of linkage to chromosomes 10p15.3-p15.1, 14q24.3-q31.1 and 9q33.3-q34.3 in non-syndromic colorectal cancer families. *Eur J Hum Genet* 2012; **20**: 91–96.
- 33 Lange C, Laird NM: On a general class of conditional tests for family-based association studies in genetics: the asymptotic distribution, the conditional power, and optimality considerations. *Genet Epidemiol* 2002; **23**: 165–180.
- 34 Van Steen K, Lange C: PBAT: a comprehensive software package for genome-wide association analysis of complex family-based studies. *Hum Genomics* 2005; **2**: 67–69.
- 35 Horvath S, Xu X, Laird NM: The family based association test method: strategies for studying general genotype–phenotype associations. *Eur J Hum Genet* 2001; **9**: 301–306.
- 36 Chen WM, Manichaikul A, Rich SS: A generalized family-based association test for dichotomous traits. *Am J Hum Genet* 2009; **85**: 364–376.
- 37 Mann JR, DuBois RN: Cyclooxygenase-2 and gastrointestinal cancer. *Cancer J* 2004; **10**: 145–152.
- 38 Lange C, Blacker D, Laird NM: Family-based association tests for survival and times-to-onset analysis. *Stat Med* 2004; **23**: 179–189.
- 39 Kato I, Toniolo P, Akhmedkhanov A, Koenig KL, Shore R, Zeleniuch-Jacquotte A: Prospective study of factors influencing the onset of natural menopause. *J Clin Epidemiol* 1998; **51**: 1271–1276.
- 40 Koo JH, Jalaludin B, Wong SKC, Kneebone A, Connor SJ, Leong RWL: Improved survival in young women with colorectal cancer. *Am J Gastroenterol* 2008; **103**: 1488–1495.
- 41 Chlebowski RT, Wactawski-Wende J, Ritenbaugh C *et al*: Estrogen plus progestin and colorectal cancer in postmenopausal women. *N Engl J Med* 2004; **350**: 991–1004.
- 42 Campbell-Thompson M, Lynch IJ, Bhardwaj B: Expression of estrogen receptor (ER) subtypes and ER β isoforms in colon cancer. *Cancer Res* 2001; **61**: 632–640.
- 43 Konstantinopoulos PA, Kominea A, Vondros G *et al*: Oestrogen receptor beta (ER β) is abundantly expressed in normal colonic mucosa, but declines in colon adenocarcinoma paralleling the tumour's dedifferentiation. *Eur J Cancer* 2003; **39**: 1251–1258.
- 44 Szczeklik W, Sanak M, Szczeklik A: Functional effects and gender association of COX-2 gene polymorphism G(-765)C in bronchial asthma. *J Allergy Clin Immunol* 2004; **114**: 248–253.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)