COLON CANCER

Analysis of candidate modifier loci for the severity of colonic familial adenomatous polyposis, with evidence for the importance of the N-acetyl transferases

M D Crabtree, C Fletcher, M Churchman, S V Hodgson, K Neale, R K S Phillips, I P M Tomlinson

Gut 2004;**53**:271–276. doi: 10.1136/gut.2003.015586

See end of article for authors' affiliations

Correspondence to: Dr I P M Tomlinson, Molecular and Population Genetics Laboratory, London Institute, Cancer Research UK, 44 Lincoln's Inn Fields, London WC2A 3PX, UK; ian.tomlinson@cancer.org.uk

Accepted for publication 30 September 2003

Background: We have recently shown that the severity of human colonic familial adenomatous polyposis (FAP) varies in a manner consistent with the action of modifier genes. These modifier genes may harbour common alleles which increase the risk of colorectal cancer (CRC) in the general population. Analyses have suggested several common polymorphisms as risk alleles for CRC.
Methods: We determined the association between the severity of colonic FAP (151 patients) and polymorphisms in *MTHFR*, *NAT1*, *NAT2*, *GSTM*, *GSTT*, cyclin D1, E-cadherin, and *APC*. All of these loci have been suggested as influencing the risk of CRC. Colonic FAP severity was quantitated as the number of polyps per colectomy specimen, standardised for colon size. We analysed the relationship between disease severity and genotype at the polymorphic site, making allowance for the position of the germline *APC* mutation.
Results: We identified significant associations between more severe disease and the absence of the *NAT1*10* genotype in the whole group of patients. In a subset of patients with germline mutations in the so-called ''mutation cluster region'', there was an association between more severe disease and the presence of *NAT2*fast* alleles. In the whole patient set, a relatively strong association existed between more severe disease and possession of both the *NAT1*non-10* and *NAT2*fast* genotypes. There was weak evidence for an association between the *APCT1493C* allele and more severe disease in the whole patient group. No consistent association with disease severity was found for the other polymorphisms.
Conclusion: The severity of colonic FAP may be modified by alleles at the *NAT1* and/or *NAT2* loci. The identity of any functional variation remains unknown as *NAT1*10* appears to be non-functional and there is linkage disequilibrium between alleles at multiple sites within these loci which are adjacent on chromosome 8p22. While evidence from this study cannot be conclusive, our data suggest that *NAT1* and *NAT2* variants may explain an approximately twofold increase in polyp number in the FAP colon.

Familial adenomatous polyposis (FAP) is a rare condition caused by a germline mutation in the *APC* gene.[1] FAP is characterised by the development of hundreds to thousands of adenomatous colonic polyps. If untreated, one or more polyps will develop into colorectal cancer, often by the fourth decade of life. The risk of cancer is related to the number of colonic adenomas.[2] The number of adenomas which an individual develops is partly a function of the position of the germline mutation, but this cannot explain the phenomenon of intrafamilial variation.[3 4] Recently, we have shown that the pattern of intrafamilial variation in colonic FAP severity is consistent with the action of modifier genes.[5 6] These observations support those made in animal models where modifier loci modulate the phenotypic severity of the *Min* mouse,[7 8] the murine equivalent of FAP.

Modifier genes of FAP are likely to be common polymorphisms. These polymorphisms are excellent candidate low penetrance alleles for differential susceptibility to colorectal cancer (CRC) in the general population.[9] CRC is known to have a heritable component[10] and many plausible risk alleles have been suggested. There are various potential strategies for their identification, including not only direct analysis of association with disease in the human population, but also complementary studies such as identification of modifier genes. Candidate FAP modifiers are potentially of several types, including carcinogen metabolism genes, DNA repair loci, tumour suppressors and oncogenes, Wnt pathway genes, and many other genes with putative roles in CRC.[11]

Metabolism of carcinogens and other chemicals foreign to the body is governed by a complex set of enzymatic reactions. Environmental carcinogens directly or indirectly damage DNA; the process is complex and consists of three elements, bioactivation, detoxification, and chemical modification. Bioactivation and detoxification are regulated by phase 1 and phase 2 enzymatic reactions. DNA damage probably occurs by the formation of adducts which then stabilise slipped mutagenic intermediates that can occur during DNA replication. The genes examined in this study are briefly outlined below. A more comprehensive review has recently been provided by Houlston and Tomlinson.[11]

N-acetyltransferases (*NAT1* and *NAT2*) are involved in phase 2 reactions which metabolise xenobiotic compounds. Two types of phenotypes are recognised for each protein—fast and slow. Slow metabolisers (for example, *NAT1*14* and *NAT1*15*) have a V_{max} approximately 50% of that of rapid acetylators.[12 13] *NAT1* and *NAT2* lie close together on chromosome 8p22. The most common variant at *NAT1* is termed

Abbreviations: FAP, familial adenomatous polyposis; CRC, colorectal cancer; NAT, N-acetyltransferase; SNP, single nucleotide polymorphism; HNPCC, hereditary non-polyposis colon cancers; GST, glutathione S-transferase; MTHF, 5,10-methylene tetrahydrofolate; MTHFR, methylene tetrahydrofolate reductase; CCND1, cyclin D1; CDH1, cellular dose of E-cadherin; MCR, mutation cluster region; PCR, polymerase chain reaction; IQR, interquartile range

Table 1 Polymorphisms tested

Polymorphism	dbSNP	Allele/phenotype frequencies	Method
MTHFR C677T	rs1801133	C: 0.76; T: 0.24	Frosst[36]
MTHFR A1298C	rs1801131	A: 0.78; C: 0.22	van der Put[26]
*NAT1**10	rs1057126 rs15561	Homo-/heterozygote: 0.43; all other genotypes: 0.57	Bell[15]
NAT2	rs1799929 rs1799930 rs1208 rs1799931	Fast acetylator allele hom/het: 0.56; slow acetylator alleles: 0.44	Smith[18]
*GSTT1**null	n/a	Null homozygote: 0.31; all other genotypes: 0.69	Pemble[37]
*GSTM1**null	n/a	Null homozygote: 0.53; all other genotypes: 0.47	Fryer[38]
CCDN1 870G/A	rs603965	G: 0.58; A: 0.42	Porter[28]
CDH1–160C/A	rs16260	C: 0.78; A: 0.22	Porter[28]
APC promoter	rs2019720	G: 0.51; A: 0.49	This study
APC V1822D	rs459552	T: 0.82; A: 0.18	This study
APC G1678A	rs42427	G: 0.11; A: 0.89	This study
APC T1493C	rs41115	T: 0.85; C: 0.15	This study

See http://www.ncbi.nlm.nih.gov/SNP/ for dbSNP details.
MTHFR, methylene tetrahydrofolate reductase; *NAT*, N-acetyltransferase; *GST*, glutathione S-transferase; *CCND1*, cyclin D1; *CDH1*, cellular dose of E-cadherin; SNP, single nucleotide polymorphism.

*NAT1*10*, which is primarily represented by two non-coding single nucleotide polymorphisms (SNPs) (table 1). It is currently thought unlikely that the *NAT1*10* alleles themselves encode functionally distinct proteins but they may be in linkage disequilibrium with fast allele functional variation.[12–14] Although not all studies are consistent, the *NAT1**10 polymorphism has been associated with a 1.92-fold increase in the risk of CRC[15 16] and with a lower age of onset of hereditary non-polyposis colon cancers (HNPCCs).[17]

For *NAT2*, humans can be divided into slow and fast acetylator populations on the basis of their isoniazid phenotype. In general, the fast phenotype corresponds to the presence of the *NAT2*4* allele, either as a heterozygote or homozygote. *NAT2*4* is defined (table 1) by a haplotype comprising specific alleles at one silent (codon 161) and three coding polymorphisms (codons 197, 268, and 286) within *NAT2*.[18] Slow acetylator status may be protective against the onset of colon cancer in patients with HNPCC,[19] although not all studies have shown this effect.[20] In a meta-analysis of low penetrance susceptibility alleles, the risk of CRC was raised in the rapid acetylator group,[11] although this effect did not reach statistical significance overall.

Glutathione S-transferases (GSTs) comprise a family of phase 2 enzymes which catalyse the reaction of glutathione with organic compounds. There is substantial (sixfold) interindividual variation in the activity of these enzymes in red blood cells.[21] The *GSTM1* null allele is a putative cancer risk allele (table 1). It has a population frequency of approximately 0.7.[22 23] *GSTM1* null allele homozygosity has been associated with a borderline increased risk of CRC.[11] Some studies have found that DNA adduct levels are higher in *GSTM1* null homozygotes if they are smokers, suggesting that gene-environment interactions are important.[24] The *GSTT1* null allele polymorphism (table 1) is responsible for variation in metabolism of mono- and dihalomethane and other similar molecules such as alkyl halides, dichloromethane, and ethylene oxide. Monohalomethanes occur naturally, but compounds such as dichloromethane and ethylene oxide are important industrial chemicals used as methylating agents, solvents, and pesticides. Sixty per cent of the human population can conjugate the above chemicals but the remainder cannot do so as they are *GSTT1* null homozygotes. In a meta-analysis, two studies showed a small non-significant adverse risk profile associated with the *GSTT1* null allele.[11]

Folate metabolism abnormalities have been implicated in the aetiology of human neural tube defects, vascular disease, and the development of malignancy.[25] Methylene tetrahydrofolate reductase (*MTHFR*) encodes an enzyme that converts 5,10-methylene tetrahydrofolate (MTHF) to 5-methyl tetrahydrofolate. Defective MTHFR function renders individuals liable to hyperhomocysteinuria and mimics the effects of dietary folate insufficiency. *MTHFR* alleles have been implicated in colorectal carcinogenesis and represent some of the strongest candidates for low penetrance susceptibility alleles.[11] The two most frequent polymorphisms (table 1) are the C677T substitution that converts alanine to valine at amino acid 222 (and is associated with reduced enzyme activity) and the C1298A that converts glutamic acid to alanine at amino acid 429.[26]

Cyclin D1 (*CCND1*) is a downstream target of β-catenin. The G870A polymorphism (table 1) interferes with the splicing of *CCND1* exon 4/5, probably reducing transcript levels.[27] This variant has been shown to be more frequent in familial colorectal cancer cases than controls[28] and may modify the age of onset of colon cancer in HNPCC.[27] CyclinD1 genotype also influences tumour number in the Apc^{Min} mouse.[29] Similarly, variation in the cellular dose of E-cadherin (*CDH1*) has caused variation in the severity of intestinal polyposis in animal models.[30] Downregulation of E-cadherin is associated with some types of intestinal, breast, bladder, and lung cancer. The *CDH1* A−160C promoter polymorphism (table 1) may alter expression of E-cadherin and is therefore a plausible candidate modifier allele for FAP.[31]

Rare *APC* variants (I1307K and E1317Q) have been associated with a raised colon cancer risk but these alleles are too infrequent to account for much of the population risk variation for "sporadic" colon cancer. However, our study of intrafamilial colonic FAP severity indicated that the severity correlation was higher between sibling pairs than parent-offspring pairs.[6] This raised the possibility that the ostensibly "wild-type" *APC* allele could be influencing disease severity. There are two relatively frequent non-silent *APC* polymorphisms in Western populations that might be responsible for the observation (table 1). These are *APC*D1822V[32] and a polymorphism (rs2019720) near the untranslated 5′ part of the gene (K Heinimann, personal communication); either could conceivably alter gene expression or function. Other non-synonymous *APC* polymorphisms include G1493A and G1678A.

We have examined the relationship between colonic FAP severity and a variety of candidate polymorphisms in the above genes (table 1).

METHODS

A total of 151 patients from 51 families were ascertained via the St Mark's Polyposis Registry, London, UK. DNA samples and clinical details were obtained from all subjects. This sample size provided >90% power to detect at $p = 0.05$ a nominal threefold increased risk of severe disease conferred by an allele at a frequency of 0.5. All patients were from pedigrees with established classical FAP—in that one or more family members had more than 100 adenomas—with autosomal dominant inheritance of disease. DNA came from one of several sources: peripheral blood; established cell line; or fixed normal tissue. Most patients had known germline *APC* mutations, between codons 170 and 1464.

The severity of colonic polyposis was assessed, as described previously.[5 6] In brief, we determined the number of adenomatous polyps in the resected colectomy specimen, after correction for length and assuming that the standard colon is 100 cm long, to give the standard count (actual count×100/patient's colon length). We did not correct for the effects of age, as we have previously shown that age is not a major source of variation in severity for patients who undergo colectomy in early adult life.[5 6]

In order to correct for effects of the germline mutation on polyp number, patients were placed into groups derived from work previously published which demonstrates that germline mutations in the *APC* ''mutation cluster region'' (MCR) tend to produce a more severe phenotype than mutations in other regions.[6 33] In practice, we used three main groupings: MCR (codons 1250–1400); MAIN (not within MCR); and UNK (germline *APC* mutation unknown).

A variety of established methods were employed to genotype most of the candidate modifier polymorphisms described above, alongside both positive control samples of known genotype and negative controls. In general, only the most common polymorphisms within each gene were typed. The references supplying details of the methods used for these polymorphisms and the nature of the variation typed are given in table 1. For the *APC* polymorphisms, we set up assays using Pyrosequencing (Uppsala, Sweden). Polymerase chain reaction (PCR) amplification was performed as recommended by the manufacturer using one biotinylated oligonucleotide, standard reagents, and cycling conditions of: 95℃ five minutes ×1; 95℃ 15 seconds, 53℃ 30 seconds, 72℃ 25 seconds ×50; 72℃ five minutes ×1. The reaction product was immobilised on streptavidin-coated Dynabeads and denatured in 0.5 M NaOH, followed by sequencing using a third oligonucleotide according to the manufacturer's standard protocol. The following oligonucleotides were used for the original PCR (table 1): for rs2019720, GTGAACAGGGTGGCAAACAG and GGCCTAACAGAGGGAG AAAAA; for rs459552, CAGACAACAAAGATTCAAAGAAACA and TCAAAAGCAAAACTTCCTCTG; for rs41115, TTCTGCCTTCTGTAGGAATGG and TAACAATCGAATCCCC TCCA; and for rs42427, AATGCTGCAGTTCAGAGGGT and CACTCAGGCTGGATGAACAA. The sequencing oligonucleotides used were, respectively, GAAAAATAATTCCAAG AACTAACAAAGACACTGGT, GAGAAGGAGTTAGAGGAGG, and ACTTTATTACATTTTGCCAC. Haplotypes were reconstructed for the *APC* polymorphisms wherever possible, partly to determine whether they were present on the germline mutant or wild-type background. None of our patients reported that they were of Ashkenazi origin and *APC* I1307K was therefore not typed.

Associations between genotypes and standard polyp count were performed, grouping genotypes at each locus, where appropriate, according to their frequencies. All statistical computations were performed using STATA 7.0. The severity data were analysed both in the whole data set and after grouping by position of the germline *APC* mutation. Our approach was to use non-parametric tests because standard polyp counts were not normally distributed in the MAIN ($n = 98$) and UNK ($n = 36$) groups, although counts in the smaller MCR group ($n = 18$) were consistent with a normally distributed variable (table 2 and Shapiro-Wilk test, details not shown). We used a threshold of $p = 0.05$ to indicate statistical significance, our reasoning being that it was more important to tolerate possible type II errors than type I errors as replication of our findings in further studies would be necessary whatever our results. Candidate polymorphisms were considered initially in isolation, and then in combination, for four groups ((i) *NAT1* and *NAT2*, (ii) *MTHFR* alleles, (iii) *GSTM1* and *GSTT1*, and (iv) the *APC* polymorphisms) because of close linkage between polymorphic sites at these loci.

RESULTS

Details of patients' colorectal polyposis are summarised in table 2. The mean standard count was 1519 polyps (SD 1548, median 950, interquartile range 533–1865). Patients included in this study were representative of the whole population of FAP patients from St Mark's Hospital (details not shown). There was no evidence of association between any polymorphism genotype (including at *APC*) and the germline mutation group (all $p > 0.2$, details not shown). Disease severity did not vary significantly with sex (Kruskal-Wallis test, $\chi^2_1 = 0.107$, $p = 0.74$). As expected, there was a strong tendency for disease to be more severe in the MCR group, with a median corrected polyp count of 3727 compared with 760 for the MAIN and 1107 for the UNK patients (Kruskal-Wallis test, $\chi^2_2 = 30.8$, $p < 0.001$). Disease severities in the MAIN and UNK groups did not differ significantly from each other ($\chi^2_1 = 0.745$, $p = 0.38$). It is highly likely that the UNK patients did not have MCR mutations because the MCR part of the gene is a small region within a large exon for which mutation detection is relatively straightforward, whereas other parts of the gene may harbour cryptic changes, such as intronic variants and large deletions. Furthermore, FAP patients with germline mutations in the MCR tend to show LOH as the ''second hit'' in their colonic polyps whereas other FAP patients tend to acquire a protein truncating mutation as their ''second hit''.[34] An FAP modifier gene might therefore act differentially with respect to the site of the germline *APC* mutation. Therefore, in addition to analysis of each polymorphism with respect to disease severity in the whole patient group, analysis was also performed for the (MAIN+UNK) and MCR groups separately.

The results of testing each polymorphism for an association with disease severity in the whole patient set are shown in table 3. The *NAT1*10* allele as heretozygote/homozygote was associated with less severe colonic polyposis (median 563) than compound heterozygotes or homozygotes for *NAT1*non-10* alleles (median 928). This association was derived solely from the (MAIN+UNK) group of patients (Kruskal-Wallis test, $\chi^2_1 = 6.61$, $p = 0.01$), with no evidence of any association in the smaller MCR group. Of the other polymorphisms (table 3), only *APCT1493C* provided any evidence of an association; there was more severe disease in those carrying the C allele, an effect which just failed to reach significance when genotypes were grouped (table 3), but was nominally significant (Kruskal-Wallis test, $\chi^2_2 = 6.00$, $p = 0.05$) when all three genotypes were tested separately, owing to the presence of severe disease in the three patients with the C/C genotype.

Table 2 Disease severity. Colonic severity is shown as the standard count (that is, the colectomy polyp count corrected for colon specimen size alone)

Mutation group	n	Mean	SD	Median	Centile 25/75	Min	Max
UNK	35	1358	1135	1107	435/1918	86	4639
MAIN	98	1133	1098	760	491/1344	66	6367
MCR	18	3934	2147	3727	2184/5000	730	7979

When the (MAIN+UNK) and MCR groups were tested separately for associations between polymorphisms and disease severity, the only significant association was that of the *NAT2* polymorphism in the MCR group. Carriers of one or two *NAT2*fast* alleles in the MCR group had a median standard polyp count of 4884 (interquartile range (IQR) 4293–6489; n = 4) compared with 1971 (IQR 1020–3198; n = 7) in those with no *NAT2*fast* allele (Kruskal-Wallis test, $\chi^2_2 = 4.32$, $p = 0.038$). *NAT2*fast* carriers did have more severe disease in the (MAIN+UNK) group also (median 900 for *NAT2*fast v* 687 for *NAT2*slow*) but this difference was not significant (Kruskal-Wallis test, $\chi^2_1 = 0.654$, $p = 0.42$). Carriers of two *GSTM1*null* alleles in the MCR group tended to have less severe disease (median 2564 (IQR 1600–3818); n = 9) than those with other *GSTM1* genotypes (median 6489 (IQR 5000–7978); n = 2) but this was of only borderline significance (Kruskal-Wallis test, $\chi^2_1 = 3.56$, $p = 0.059$). All polymorphisms which demonstrated associations with disease in the (MAIN+UNK) group alone also did so in the whole set of patients.

We tested the effects of combinations of genotypes in different ways, depending on the loci concerned: for *NAT1* and *NAT2*, we tested the combination of absence of *NAT1*10* and possession of *NAT2*fast*; for *MTHFR*, we used the *hapipf* command of STATA to predict haplotypes; for *GSTM1* and *GSTT1*, we tested the combination of null alleles at each locus; and for *APC*, we directly determined ''wild-type'' haplotypes using the pedigrees. There was no detectable association between genotype and disease severity for *MTHFR*, *GSTM1-GSTT1*, and *APC* (details not shown). For *NAT1–NAT2*, however, more severe disease was associated with combined absence of the *NAT1*10* genotype and presence of the *NAT2*fast* genotype. This association was present both in the entire patient set (median 1167 for *NAT1*non-10-NAT2*fast v* 687 for other genotypes; Kruskal-Wallis test, $\chi^2_1 = 4.34$, $p = 0.037$) and particularly in the MAIN/UNK group (median 1167 for *NAT1*non-10; NAT2*fast v* 587 for other genotypes; Kruskal-Wallis test, $\chi^2_1 = 10.32$, $p = 0.0013$).

DISCUSSION

We have found evidence to suggest that the severity of colonic polyposis in FAP is associated with variation at the *NAT1* and/or *NAT2* loci. The *NAT1*10* allele was associated with less severe disease in the entire patient set, and the *NAT2*fast* type was associated with more severe disease in the subset of patients with a germline mutation in the MCR. *NAT1*10* has previously been associated with an increased risk of colon cancer,[15] apparently in contradistinction to our data; but *NAT*10* is unlikely to represent functional variation[13 14] and any association with disease is likely to come from linkage disequilibrium with functional variants. There is

Table 3 Associations between disease severity and genotype at each polymorphism in the entire patient set

Polymorphism	Genotype (n)	Standard count (median (IQR))	Significance
MTHFR C677T	*C/C* (54)	973 (526–1465)	$\chi^2_1 = 0.164$
	C/T, T/T (40)	811 (490–2078)	p = 0.66
MTHFR A1298C	*A/A* (55)	892 (490–1676)	$\chi^2_1 = 0.268$
	A/C, C/C (39)	865 (446–1465)	p = 0.64
NAT1	*NAT1*10* (30)	563 (320–1295)	$\chi^2_1 = 4.42$
	Other (48)	928 (641–1638)	p = 0.039
NAT2	*Fast* (52)	955 (482–1786)	$\chi^2_1 = 0.165$
	Slow (41)	825 (490–1367)	p = 0.69
GSTM1	*Null* (50)	920 (561–1511)	$\chi^2_1 = 0.347$
	Non-null (43)	748 (435–1892)	p = 0.56
GSTT1	*Null* (35)	1020 (561–1679)	$\chi^2_1 = 0.625$
	Non-null (60)	811 (445–1630)	p = 0.43
CCDN1 870G/A	*G/G* (32)	845 (375–1702)	$\chi^2_2 = 1.46$
	G/A (72)	926 (541–1922)	p = 0.48
	A/A (15)	748 (368–1591)	
CDH1–160C/A	*C/C* (70)	989 (490–1926)	$\chi^2_1 = 1.57$
	C/A, A/A (40)	706 (445–1639)	p = 0.46
APC promoter	*G/G* (21)	1088 (490–2430)	$\chi^2_2 = 1.27$
	G/A (46)	713 (463–1295)	p = 0.53
	A/A (26)	1025 (646–1591)	
APC V1822D	*T/T* (66)	896 (526–1808)	$\chi^2_1 = 2.32$
	T/A, A/A (33)	947 (539–1591)	p = 0.13
APC G1678A	*A/A* (68)	924 (535–1630)	$\chi^2_1 = 2.08$
	A/G, G/G (20)	927 (593–2078)	p = 0.35
APC T1493C	*T/T* (71)	892 (463–1659)	$\chi^2_1 = 3.22$
	T/C, C/C (28)	1077 (604–2078)	p = 0.073

MTHFR, methylene tetrahydrofolate reductase; *NAT*, N-acetyltransferase; *GST*, glutathione S-transferase; *CCND1*, cyclin D1; *CDH1*, cellular dose of E-cadherin; IQR, interquartile range.
The Kruskal-Wallis test was used to test for associations between the genotypes shown and the standard polyp count. n, number of patients successfully analysed for each genotype.

some existing evidence to suggest that *NAT2*fast* alleles are functionally associated with an increased risk of bowel cancer,[19] although again, linkage disequilibrium cannot be excluded. To further complicate matters, *NAT1* and *NAT2* lie within 0.2 Mb on chromosome 8p22 and there is evidence that at least some alleles at each locus are in moderate linkage disequilibrium with each other.[35] We found that the combined absence of *NAT1*non-10* and the presence of *NAT2*fast* were strongly associated with severe disease.

Given that variation in FAP polyp number appears to be related to tumour initiation rather than progression,[5] a differential effect of the *NAT* phenotype might influence the early stages of tumorigenesis in the FAP colon, perhaps even at the stage of the ''second hit'' at *APC*. If this is the case, then carcinogen metabolism is likely to be important in early human growth and development. It is even possible that carcinogens from the maternal diet impact on the frequency and type of the ''second hit'' in utero.

The possible association of the *GSTM1* non-null phenotype with more severe colonic FAP in the MCR patients appears unlikely to result from a true biological effect. The possible association is somewhat puzzling on theoretical grounds and on the basis of previous studies, from which we would predict the opposite to our findings. We note, moreover, that—unlike *NAT1* and *NAT2*—there was no evidence for the *GSTM1* severity association in the largest (MAIN) group of patients. Perhaps given that the five *GSTM* loci lie within 100 kb on chromosome 1p13.3, there is a true association between another *GSTM* allele and disease severity.

Similarly, while we have previously suggested that variation in the *APC* gene itself might modify the FAP phenotype, we are cautious about the finding of a possible association between the *APC*1493C allele and more severe disease. This variation is unlikely to have functional consequences but may be in linkage disequilibrium with other variation, for example in the *APC* promoter (either in *cis* or in *trans* with the germline *APC* mutant). Haplotype reconstruction from pedigrees did not however show that the C allele was consistently associated with the mutant or wild-type *APC* allele and we found no evidence for a specific *APC* haplotype associated with disease severity. It is possible that the potential association between *APC*1493C and disease severity results in part from the fact that we decided to include different members of each family as providing independent data for this study. The justification for this relies on the following: (i) balancing potential errors from the assumption of independence against the loss of power; (ii) our analysis of common polymorphisms which are expected to segregate within families; (iii) relatively small number of individuals per family in our study (median 3); and (iv) allowance for the effects of different germline *APC* mutations by subdividing into MAIN, UNK, and MCR groups.

Our findings on the other polymorphisms (*MTHFR*, *GSTT1*, E-cadherin, and cyclin D1) provide no convincing evidence for an effect on disease severity. In the case of the *MTHFR* C677T polymorphism, in particular, this is an important negative result. Studies have consistently shown a weak association of the T allele homozygote with reduced colon cancer risk (summarised by Houlston and Tomlinson[11]) but no effect on adenoma risk has been detected, in accord with our results.

We conclude that variation at the *NAT1* and/or *NAT2* loci is a possible modifier of the number of colorectal polyps in patients with FAP. The nature of the variation remains unknown. Evidence from this study is not conclusive but our data suggest that *NAT1* variants may explain an approximately twofold increase in polyp number in the FAP colon (table 3). The same or different modifier genes may be involved in determining the severity of extracolonic disease, such as desmoids and upper gastrointestinal tumours. It is however evident that, even if the effects of *NAT1* and/or *NAT2* are confirmed by other studies, there may still be unexplained variation in FAP. The possibility remains that variation in FAP severity is a polygenic trait, its study requiring analysis of large numbers of patients at multiple polymorphic loci.

ACKNOWLEDGEMENTS

We are grateful to the patients involved in the study, to the Equipment Park, Cancer Research UK, to staff from the Polyposis Registry and Colorectal Cancer Unit, Cancer Research UK, St Mark's Hospital, and to two referees (Tim Bishop and one anonymous). MC held a Cancer Research UK Translational Fellowship.

Authors' affiliations

M D Crabtree, I P M Tomlinson, Molecular and Population Genetics Laboratory, London Institute, Cancer Research UK, London, UK, and the Academic Institute, St Mark's Hospital, Harrow, Middlesex, UK
C Fletcher, M Churchman, Genotyping Group, Cancer Research UK, Henry Wellcome Building, Headington, Oxford, UK
K Neale, R K S Phillips, The Academic Institute, St Mark's Hospital, Harrow, Middlesex, UK
S V Hodgson, Department of Genetics, 8th Floor, Guy's Tower, Guy's Hospital, London, UK

REFERENCES

1 **Fearnhead NS**, Britton MP, Bodmer WF. The ABC of APC. *Hum Mol Genet* 2001;**10**:721–33.
2 **Debinski HS**, Love S, Spigelman AD, *et al.* Colorectal polyp counts and cancer risk in familial adenomatous polyposis. *Gastroenterology* 1996;**110**:1028–30.
3 **Paul P**, Letteboer T, Gelbert L, *et al.* Identical APC exon 15 mutations result in a variable phenotype in familial adenomatous polyposis. *Hum Mol Genet* 1993;**2**:925–31.
4 **Scott RJ**, Meldrum C, Crooks R, *et al.* Familial adenomatous polyposis: more evidence for disease diversity and genetic heterogeneity. *Gut* 2001;**48**:508–14.
5 **Crabtree MD**, Tomlinson IP, Talbot IC, *et al.* Variability in the severity of colonic disease in familial adenomatous polyposis results from differences in tumour initiation rather than progression and depends relatively little on patient age. *Gut* 2001;**49**:540–3.
6 **Crabtree MD**, Tomlinson IP, Hodgson SV, *et al.* Explaining variation in familial adenomatous polyposis: relationship between genotype and phenotype and evidence for modifier genes. *Gut* 2002;**51**:420–3.
7 **Dietrich WF**, Lander ES, Smith JS, *et al.* Genetic identification of Mom-1, a major modifier locus affecting Min-induced intestinal neoplasia in the mouse. *Cell* 1993;**75**:631–9.
8 **MacPhee M**, Chepenik KP, Liddell RA, *et al.* The secretory phospholipase A2 gene is a candidate for the Mom1 locus, a major modifier of ApcMin-induced intestinal neoplasia. *Cell* 1995;**81**:957–66.
9 **Houlston R**, Crabtree M, Phillips R, *et al.* Explaining differences in the severity of familial adenomatous polyposis and the search for modifier genes. *Gut* 2001;**48**:1–5.
10 **Lichtenstein P**, Holm NV, Verkasalo PK, *et al.* Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 2000;**343**:78–85.
11 **Houlston RS**, Tomlinson IP. Polymorphisms and colorectal tumor risk. *Gastroenterology* 2001;**121**:282–301.
12 **Hein DW**, Grant DM, Sim E. Update on consensus arylamine N-acetyltransferase gene nomenclature. *Pharmacogenetics* 2000;**10**:291–2.
13 **Hein DW**, Doll MA, Fretland AJ, *et al.* Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms. *Cancer Epidemiol Biomarkers Prev* 2000;**9**:29–42.
14 **Hein DW**. Molecular genetics and function of NAT1 and NAT2: role in aromatic amine metabolism and carcinogenesis. *Mutat Res* 2002;**506–507**:65–77.
15 **Bell DA**, Badawi AF, Lang NP, *et al.* Polymorphism in the N-acetyltransferase 1 (NAT1) polyadenylation signal: association of NAT1*10 allele with higher N-acetylation activity in bladder and colon tissue. *Cancer Res* 1995;**55**:5226–9.
16 **Bell DA**, Stephens EA, Castranio T, *et al.* Polyadenylation polymorphism in the acetyltransferase 1 gene (NAT1) increases risk of colorectal cancer. *Cancer Res* 1995;**55**:3537–42.
17 **Moisio AL**, Sistonen P, Mecklin JP, *et al.* Genetic polymorphisms in carcinogen metabolism and their association to hereditary nonpolyposis colon cancer. *Gastroenterology* 1998;**115**:1387–94.
18 **Smith CA**, Wadelius M, Gough AC, *et al.* A simplified assay for the arylamine N-acetyltransferase 2 polymorphism validated by phenotyping with isoniazid. *J Med Genet* 1997;**34**:758–60.
19 **Heinimann K**, Scott RJ, Chappuis P, *et al.* N-acetyltransferase 2 influences cancer prevalence in hMLH1/hMSH2 mutation carriers. *Cancer Res* 1999;**59**:3038–40.

20 **Frazier ML**, O'Donnell FT, Kong S, *et al.* Age-associated risk of cancer among individuals with N-acetyltransferase 2 (NAT2) mutations and mutations in DNA mismatch repair genes. *Cancer Res* 2001;**61**:1269–71.
21 **Scott EM**, Wright RC. Variability of glutathione S-transferase of human erythrocytes. *Am J Hum Genet* 1980;**32**:115–7.
22 **Board PG**, Johnston PN, Ross VL, *et al.* Molecular genetics of the human glutathione S-transferase. *Princess Takamatsu Symp* 1990;**21**:199–211.
23 **Board P**, Coggan M, Johnston P, *et al.* Genetic heterogeneity of the human glutathione transferases: a complex of gene families. *Pharmacol Ther* 1990;**48**:357–69.
24 **Firozi PF**, Bondy ML, Sahin AA, *et al.* Aromatic DNA adducts and polymorphisms of CYP1A1, NAT2, and GSTM1 in breast cancer. *Carcinogenesis* 2002;**23**:301–6.
25 **Rosenblatt DS**, Whitehead VM. Cobalamin and folate deficiency: acquired and hereditary disorders in children. *Semin Hematol* 1999;**36**:19–34.
26 **van der Put NM**, Gabreels F, Stevens EM, *et al.* A second common mutation in the methylenetetrahydrofolate reductase gene: an additional risk factor for neural-tube defects? *Am J Hum Genet* 1998;**62**:1044–51.
27 **Kong S**, Amos CI, Luthra R, *et al.* Effects of cyclin D1 polymorphism on age of onset of hereditary nonpolyposis colorectal cancer. *Cancer Res* 2000;**60**:249–52.
28 **Porter TR**, Richards FM, Houlston RS, *et al.* Contribution of cyclin d1 (CCND1) and E-cadherin (CDH1) polymorphisms to familial and sporadic colorectal cancer. *Oncogene* 2002;**21**:1928–33.
29 **Wilding J**, Straub J, Bee J, *et al.* Cyclin D1 is not an essential target of beta-catenin signaling during intestinal tumorigenesis, but it may act as a modifier of disease severity in multiple intestinal neoplasia (Min) mice. *Cancer Res* 2002;**62**:4562–5.
30 **Smits R**, Ruiz P, Diaz-Cano S, *et al.* E-cadherin and adenomatous polyposis coli mutations are synergistic in intestinal tumor initiation in mice. *Gastroenterology* 2000;**119**:1045–53.
31 **Li LC**, Chui RM, Sasaki M, *et al.* A single nucleotide polymorphism in the E-cadherin gene promoter alters transcriptional activities. *Cancer Res* 2000;**60**:873–6.
32 **Ruiz-Ponte C**, Vega A, Conde R, *et al.* The Asp1822Val variant of the APC gene is a common polymorphism without clinical implications. *J Med Genet* 2001;**38**:E33.
33 **Nugent KP**, Phillips RK, Hodgson SV, *et al.* Phenotypic expression in familial adenomatous polyposis: partial prediction by mutation analysis. *Gut* 1994;**35**:1622–3.
34 **Lamlum H**, Ilyas M, Rowan A, *et al.* The type of somatic mutation at APC in familial adenomatous polyposis is determined by the site of the germline mutation: a new facet to Knudson's 'two-hit' hypothesis. *Nat Med* 1999;**5**:1071–5.
35 **Smelt VA**, Mardon HJ, Sim E. Placental expression of arylamine N-acetyltransferases: evidence for linkage disequilibrium between NAT1*10 and NAT2*4 alleles of the two human arylamine N-acetyltransferase loci NAT1 and NAT2. *Pharmacol Toxicol* 1998;**83**:149–57.
36 **Frosst P**, Blom HJ, Milos R, *et al.* A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nat Genet* 1995;**10**:111–13.
37 **Pemble S**, Schroeder KR, Spencer SR, *et al.* Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism. *Biochem J* 1994;**300**(Pt 1):271–6.
38 **Fryer AA**, Zhao L, Alldersea J, *et al.* The glutathione S-transferases: polymerase chain reaction studies on the frequency of the GSTM1 0 genotype in patients with pituitary adenomas. *Carcinogenesis* 1993;**14**:563–6.