

COLORECTAL CANCER

Explaining variation in familial adenomatous polyposis: relationship between genotype and phenotype and evidence for modifier genes

M D Crabtree, I P M Tomlinson, S V Hodgson, K Neale, R K S Phillips, R S Houlston

Gut 2002;51:420–423

See end of article for authors' affiliations

Correspondence to: M D Crabtree, Molecular and Population Genetics Laboratory, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, UK; m.crabtree@icrf.icnet.uk

Accepted for publication 9 January 2002

Background: Familial adenomatous polyposis (FAP) is characterised by variable phenotypic expression. Part of this is attributable to a relationship between *APC* genotype and phenotype but there remains significant intrafamilial variation. In the *Min* mouse model of FAP, differences in the severity of gastrointestinal polyposis result from the action of modifier genes.

Aims: To determine whether phenotypic variation in human FAP has an inherited component consistent with the action of modifier genes.

Method: We systematically examined polyp numbers in colectomy specimens from patients with classical FAP. Variation both between and within families was analysed. Formal modelling of the segregation of disease severity in families was performed

Results: There was strong evidence for a relationship between site of mutation and the number of colorectal polyps, with germline mutations in the "cluster region" causing the most severe disease and those with mutations between codons 1020 and 1169 having the mildest disease. In addition to this genotype-phenotype relationship, we found evidence for non-*APC* linked genetic modifiers of disease expression. First degree relatives had more similar polyp counts than more distant relatives. Formal modelling of the segregation of disease severity in families revealed further evidence for the action of modifier genes, with a best fit to a mixed model of inheritance.

Conclusion: Our data provide good evidence to support the hypothesis that modifier genes influence the severity of FAP in humans.

Familial adenomatous polyposis (FAP) is a disease characterised by multiple colorectal adenomas, specific extracolonic features, and a greatly increased risk of gastrointestinal cancer. Most of the clinical features of FAP are characterised by variable expression. For example, the number of adenomatous polyps in the large bowel can vary considerably, independently of age. Furthermore, many extracolonic features, such as intra-abdominal desmoids, are seen in only a minority of cases. Some of the variable expression of FAP reflects the influence of different germline *APC* mutations.^{1–3} It has been shown, for example, that mutations in exon 9 and towards the 5' and 3' ends of the *APC* gene are associated with an attenuated phenotype.^{3–6} With respect to classical FAP, the greatest numbers of colorectal polyps are probably produced by germline mutations close to codon 1309.^{1–7} However, apart from these specific sites within the *APC* gene, the dependence of polyp number on genotype is unknown.

Variation observed within families or groups of individuals with the same germline mutation suggests that factors other than *APC* genotype can affect disease expression in FAP. Part of the unexplained phenotypic variation may result from the action of modifier genes not linked to *APC*. Studies of the *Min* (multiple intestinal neoplasia) mouse model of FAP have shown that the numbers of gastrointestinal polyps vary among strains.^{8–9} Linkage studies have mapped modifier loci to mouse chromosome 4, with secretory phospholipase A₂ a prime candidate for a modifier of *Min*.¹⁰ Linkage studies in humans have however revealed only weak evidence for a modifier gene on chromosome 1p35-p36, the region syntenic to that containing *Mom1*.^{11–12} No functional variants in secretory phospholipase A₂ have been identified in humans.^{13–15} Thus there is no direct evidence that the unexplained variation in FAP severity results from inherited

factors rather than from alternative mechanisms such as different environments or simply by chance.

Our aims in this study on variation in the severity of colonic FAP, as assessed by number of polyps at prophylactic colectomy, were twofold. Firstly, we analysed genotype-phenotype associations in classical FAP according to functional domains of the *APC* protein. Secondly, we examined the evidence that modifier loci other than *APC* act to determine the severity of colonic polyposis in humans, and investigated which genetic model, if any, is most parsimonious with respect to the observed pattern of familial aggregation.

METHODS

Data collection

In total, 221 patients (124 male, 97 female) with classical FAP (>100 colorectal adenomas) who had undergone prophylactic colectomies were entered into this study, having been systematically ascertained through the St Mark's Hospital Polyposis Registry, Harrow, UK. The germline *APC* mutation was known for 175 patients who were therefore employed for the analysis of genotype-phenotype associations. A total of 166 individuals from 55 families were available for intrafamilial analysis of modifying genes; in 128 of these patients, an *APC* mutation had been identified. Median age of patients at colectomy was 21.1 years and 45% were aged 14–21 years at the time of colectomy (interquartile range 17.8–29.7 years). Polyp counts on colectomy specimens were performed routinely, as described previously,¹⁶ and corrected for differences in colon

Abbreviations: FAP, familial adenomatous polyposis; MCR, mutation cluster region; *Min*, multiple intestinal neoplasia.

length. Although adenoma number appears to depend on patient age—for example, very young children with *APC* mutations rarely have macroscopic polyps—we have previously shown that age is unlikely to be a major determinant of severity in the patients analysed in this study because polyp counts change little over time once patients are in their teens and older.¹⁶ Nevertheless, we reassessed the influence of age on the number of polyps using regression analysis in the group of patients studied here, some of whom had been studied in the previous analysis.

Relationship between genotype and phenotype

Germline *APC* mutations were denoted by their nucleotide and/or codon position using the *APC* cDNA reference sequence (Genbank NM_000038). To analyse the relationship between genotype and phenotype, *APC* mutations were grouped according to their position within the functional domains of the gene (fig 1). The presumption was that truncating mutations in the same region of the gene are likely to result in similar phenotypic effects. The domains defined were: (i) pre-armadillo region (codons 168–453); (ii) armadillo repeat region plus (codons 454–1019); (iii) β -catenin binding region (codons 1020–1168); (iv) post- β -catenin binding region (codons 1169–1250); (v) central mutation cluster region (MCR, codons 1250–1400); and (vi) post-MCR (codons 1400–1580). Affected members of the same family were all assumed to have an identical *APC* mutation, whether or not the mutation had been established within all affected family members. Data were log transformed on the basis of the observed distribution, thereby improving the normality of the different groups and reducing the differences in regional variances (details not shown). Differences in the severity of polyposis (number of adenomas at colectomy) between regions were assessed using F and *t* tests. In order to adjust for multiple testing, the Bonferonni correction was applied to significance levels.

Evidence for modifying genes

In general, closer relatives from a FAP kindred would tend to share alleles at a modifier locus and thus have more similar phenotypes. To examine intrafamilial effects independent of germline *APC* mutation, we first computed correlation coefficients between different types of relative pairs within families (siblings, parent-offspring, and second degree relatives) using log transformed data which had been corrected for germline mutations. All standard statistical analyses were undertaken using the program STATA (version 7.0; Stata Corporation, 702 University Drive East, College Station, Texas 77840, USA). To assess the evidence for a major gene component against polygenic effects, we performed a segregation analysis under a mixed model¹⁷ using the program PAP, version 4.0.¹⁸ Evidently, as only individuals with FAP could be scored for number of polyps, unaffected individuals were treated as unknown. Significance tests were based on standard likelihood ratio tests (that is, to test competing hypotheses, minus twice the log likelihood calculated under the general model was subtracted from the likelihood when one or more parameters were held constant, the difference being distributed as a χ^2 with a number of degrees of freedom equal to the number of free parameters).

RESULTS

There was no significant relationship between age and number of adenomas ($r^2=0.09$, $p>0.3$) after patients were stratified by mutation, implying that polyp number changes little around the time that colectomies are undertaken. These data are therefore in agreement with our previous analysis using paired data which showed no significant effect of age on polyp count in patients having prophylactic colectomy.¹⁶

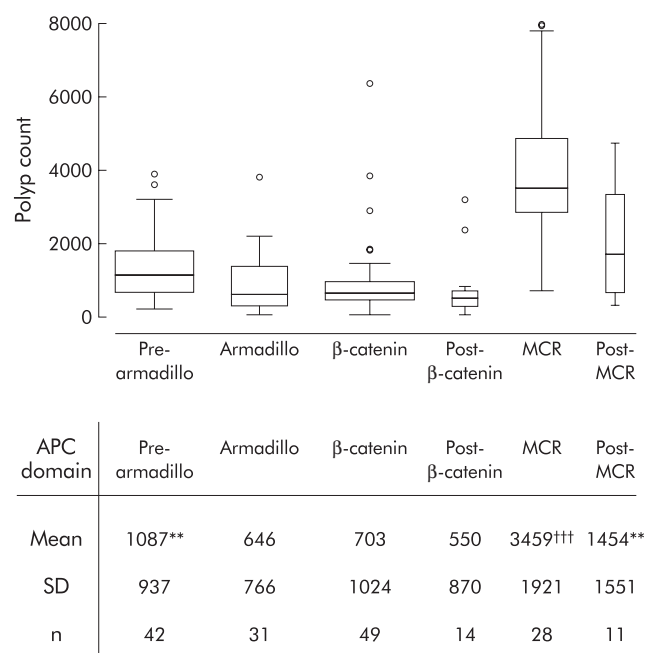


Figure 1 Box and whisker plot and aligned table show the relationship between severity of familial adenomatous polyposis and *APC* protein domain. The horizontal lines correspond to the 5th, 25th, 50th, 75th, and 95th centiles. The *APC* domains are arranged in their order on the gene 5' to 3', left to right. The data shown are untransformed—that is, raw polyp counts—while calculations used in this paper were on log transformed data values. Summary statistics are shown in the table. Three severity groups were distinguished using multiple unpaired *t* tests: severe (mutation cluster region (MCR) region), moderate (pre-armadillo and post-MCR), and mild (all other groups). ** $p<0.01$ compared with MCR, armadillo, β -catenin, and post- β -catenin (unpaired *t* test); ††† $p<0.001$ compared with all other regions (unpaired *t* test).

Therefore, for the purpose of this specific analysis, age correction was not applied. Figure 1 shows the severity of colonic polyposis according to mutations in each of the six functional domains of *APC*. The most severe disease was associated with mutations within the MCR (mean polyp count 3459). Mutations within the pre-armadillo region and 3' to the MCR were associated with moderate polyposis (mean polyp counts 1087 (937.4) and 1454 (1551), respectively), while mutations in the remaining domains were associated with mild colonic polyposis (armadillo(+), mean 646 (766.3); β -catenin, mean 703 (1024); post- β -catenin, mean 550 (870.6)).

Twenty two of the affected individuals harboured mutations within the MCR of *APC*. Fourteen of these were carriers of the 1309 del5bp mutation. There was no difference in the number of colorectal adenomas in these patients compared with those who possessed other MCR mutations ($p=0.66$). Similarly, there was no difference between the number of adenomas in the 29 patients with the common mutation at codon 1061 in the β -catenin binding domain compared with individuals with other mutations in this domain ($p=0.98$).

To compute correlations between different groups of relatives and combine data across the different *APC* domain groups, trait values were Z transformed:

$$z_i = (x_i - \mu_g) / s_g$$

where x_i is log transformed polyp count; μ_g is the mean for the *APC* domain group; and s_g is the standard deviate for the *APC* domain group.

There was a highly significant correlation ($r=0.42$) in the severity of disease in sibling pairs (table 1). The correlation between parent-offspring pairs ($r=0.29$) was also significant but weaker whereas the correlations between more distantly

Table 1 Pairwise correlations between relatives for familial adenomatous polyposis severity

Canonical relationship	Genetic identity by descent	n	Correlation	p Value
Sibling	50%	70	0.42	0.0002
Parental	50%	66	0.29	0.010
Avuncular	25%	33	0.26	0.115
1st cousin	12.5%	34	0.08	0.618
2nd cousin once removed	6.3%	35	0.21	0.234
2nd cousin	3.1%	28	0.28	0.157

Table 2 Segregation analysis for within family variation in familial adenomatous polyposis

Model	Gene frequency	Polygenic heritability	2LogL*
Dominant	0.54	—	12.5
Recessive	0.15	—	11.0
Polygenic	—	0.47	7.0
Mixed	0.66	0.25	20.2

*Log likelihoods are relative to the model with no familial effect.

related relatives were much weaker and non-significant (table 1). This pattern of familial correlation is consistent with a genetic component influencing FAP severity but not related to *APC* mutation.

Table 2 shows the results of segregation analysis. Single gene models fitted the data better than a simple polygenic model. Of the two single gene models, the dominant model fitted the observed data better than the recessive model. There was however evidence from the segregation analysis to suggest that a single gene genetic model is not adequate to explain the distribution of polyps in families and that a more complex model is required. Both single gene models fitted the data less well than the mixed model ($\chi^2=7.7$, comparing a dominant model with a mixed model; $\chi^2=9.2$, comparing a recessive model with the mixed model).

DISCUSSION

There are a number of possible sources of variation in the expression of the FAP phenotype. In our study of classical FAP, we have found clear evidence of a relationship between the site of the germline *APC* mutation and the severity of the colonic phenotype, as measured by the number of adenomatous polyps at prophylactic colectomy. While previous data emphasised the different mutation spectrum in classical and attenuated FAP,³ the common codon 1309 mutation was the only genotype associated with more severe colonic polyposis.¹ We have shown that severe disease is associated not only with mutations at codon 1309 but that other MCR mutations confer comparable disease. In addition, the data indicate that patients with mutations in codons 168–453 and 1400–1465 of *APC* have moderate polyposis and mutations outside these regions confer a milder phenotype.

Although there is evidence of a relationship between genotype and phenotype, there is also evidence of intrafamilial variability in FAP.¹⁹ The pattern of variation within families provides some indication as to the aetiology of this intrafamilial variation. The higher phenotypic correlations between first degree than second degree relatives would certainly be expected for effects produced by modifying genes at unlinked loci but might also result from environmental factors that are more likely to be shared among closer relatives. Determining the role of any environmental familial factor can, in principle, be addressed by examining concordance in monozygotic

twins. There were no monozygotic twins available for analysis in our study and, given the comparative rarity of FAP, accruing sufficient numbers of monozygotic twins to make this a viable proposition is unlikely. It is however difficult to envisage environmental influences such as dietary carcinogens having a profound effect on the severity of FAP as the disease manifests early in life and patients appear to differ in absolute adenoma number rather than in the rate of progression from microscopic to macroscopic lesions.¹⁶

Although both types of first degree relative pairs were more similar than all other relative pairs, we found greater phenotypic similarity between sibs than between parents and offspring. Both sibling pairs and parent-offspring pairs are first degree relatives and are therefore expected to share the same proportion of non-allelic modifier alleles. Any differences detected in correlation might result from differences in environmental sharing or temporal trends in data collection but might alternatively result from genetic variation linked to *APC*, such as an effect of the ostensibly "wild-type" *APC* allele. Affected siblings share the same wild-type allele by descent from the unaffected parent for 50% of the time in contrast with parent-offspring pairs who will more rarely do so. Functional polymorphisms of the normal *APC* allele might therefore provide a genetic explanation for a higher correlation between siblings than among parent-offspring pairs. There is evidence for such a mechanism in some mendelian diseases.^{20–21} While to date there is no direct evidence for such an effect in FAP, functional polymorphisms such as *APC1307K*²² do exist.

Since the identification of *Mom1* in the *Min* mouse,^{9–10} the existence of modifier genes for FAP has been the subject of intense speculation. We have provided evidence that the clinical phenotype in human colonic FAP is consistent with the action of modifier genes. Direct confirmation of the hypothesis will require identification of the responsible genes. In the absence of a clear model of inheritance of FAP severity in our data, it is likely that identification of FAP modifier genes will rely on testing candidate loci. These candidates include genes implicated in colorectal carcinogenesis a priori such as *COX2*, *MTHFR*, and DNA repair genes, and genes located within chromosomal regions syntenic to susceptibility loci in mouse models. Although there may be some limited benefits for clinical practice in FAP, identification of FAP modifier genes is most likely to be useful in providing insight into mechanisms of colorectal tumorigenesis, indicating possible therapeutic targets and suggesting excellent candidates for colon cancer susceptibility genes in general.

ACKNOWLEDGEMENTS

The authors acknowledge Kalpna Pindoria's help.

Authors' affiliations

M D Crabtree, Molecular and Population Genetics Laboratory, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, UK, and the Academic Institute and Polyposis Registry, St Mark's Hospital,

Watford Road, Harrow, Middlesex HA1 3UJ, UK

I P M Tomlinson, Molecular and Population Genetics Laboratory, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, UK

S V Hodgson, Department of Clinical Genetics, Guy's Hospital, St Thomas St, London SE1 9RT, UK

K Neale, R K S Phillips, Academic Institute and Polyposis Registry, St Mark's Hospital, Watford Road, Harrow, Middlesex HA1 3UJ, UK

R S Houlston, Section of Cancer Genetics, Institute of Cancer Research, Cotswold Road, Sutton SM2 5NG, UK

REFERENCES

- 1 **Nugent K**, Phillips R, Hodgson S, *et al.* Phenotypic expression in familial adenomatous polyposis: partial prediction by mutation analysis. *Gut* 1994;**35**:1622-3.
- 2 **Wu J**, Paul P, McGannon E, *et al.* APC genotype, polyp number, and surgical options in familial adenomatous polyposis. *Ann Surg* 1998;**227**:57-62.
- 3 **Spirio L**, Olschwang S, Groden J, *et al.* Alleles of the APC gene: an attenuated form of familial polyposis. *Cell* 1993;**75**:951-7.
- 4 **Samowitz WS**, Thliveris A, Spirio LN, *et al.* Alternatively spliced adenomatous polyposis coli (APC) gene transcripts that delete exons mutated in attenuated APC. *Oncogene* 1995;**55**:3732-4.
- 5 **Friedl W**, Meuschel S, Caspari R, *et al.* Attenuated familial adenomatous polyposis due to a mutation in the 3' part of the APC gene. A clue for understanding the function of the APC protein. *Hum Genet* 1996;**97**:579-84.
- 6 **Soravia C**, Berk T, Madlensky L, *et al.* Genotype-phenotype correlations in attenuated adenomatous polyposis coli. *Am J Hum Genet* 1998;**62**:1290-301.
- 7 **Giardiello F**, Krush A, Petersen G, *et al.* Phenotypic variability of familial adenomatous polyposis in 11 unrelated families with identical APC gene mutation. *Gastroenterology* 1994;**106**:1542-7.
- 8 **Su LK**, Kinzler KW, Vogelstein BV, *et al.* Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. *Science* 1992;**256**:668-70.
- 9 **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.
- 10 **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.
- 11 **Tomlinson IPM**, Neale K, Talbot IC, *et al.* A modifying locus for familial adenomatous polyposis may be present on chromosome 1p35-p36. *J Med Genet* 1996;**33**:268-73.
- 12 **Dobbie Z**, Heinemann K, Bishop D, *et al.* Identification of a modifier gene locus on chromosome 1p35-36 in familial adenomatous polyposis. *Hum Genet* 1997;**99**:653-7.
- 13 **Riggins GJ**, Markowitz S, Wilson JK, *et al.* Absence of secretory phospholipase A2 gene alterations in human colorectal cancer. *Cancer Res* 1995;**55**:5184-6.
- 14 **Dobbie Z**, Muller H, Scott R. Secretory phospholipase A2 does not appear to be associated with phenotypic variation in familial adenomatous polyposis. *Hum Genet* 1996;**98**:386-90.
- 15 **Tomlinson I**, Beck N, Neale K, *et al.* Variants at the secretory phospholipase A2 (PLA2G2A) locus: analysis of associations with familial adenomatous polyposis and sporadic colorectal tumours. *Ann Hum Genet* 1996;**60**:369-76.
- 16 **Crabtree M**, Tomlinson I, Talbot I, *et al.* Variability in the severity of colonic polyposis in FAP results from differences in tumour initiation rather than progression and depends relatively little on patient age. *Gut* 2001;**49**:540-3.
- 17 **Morton NE**, MacLean CJ. Analysis of family resemblance. 3. Complex segregation of quantitative traits. *Am J Hum Genet* 1974;**26**:489-503.
- 18 **Hastadt S**. *Pedigree analysis package*, rev. 4.0. Salt Lake City: University of Utah, 1994.
- 19 **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.
- 20 **Maugeri A**, van Driel MA, van *et al.* The 2588G→C mutation in the ABCR gene is a mild frequent founder mutation in the Western European population and allows the classification of ABCR mutations in patients with Stargardt disease. *Am J Hum Genet* 1999;**64**:1024-35.
- 21 **Gouya L**, Puy H, Lamoril J, *et al.* Inheritance in erythropoietic protoporphyria: a common wild-type ferrochelatase allelic variant with low expression accounts for clinical manifestation. *Blood* 1999;**93**:2105-10.
- 22 **Laken SJ**, Petersen GM, Gruber SB, *et al.* Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. *Nat Genet* 1997;**17**:79-83.