Genetics of inflammatory bowel disease

J Satsangi, D P Jewell, W M C Rosenberg, J I Bell

Introduction
The aetiology and pathogenesis of the chronic inflammatory bowel diseases, Crohn's disease and ulcerative colitis, remain under investigation. It has become accepted that disease occurs in a genetically susceptible subject, as a result of the interaction between antigenic stimulus and the cells of the immune system. As a preliminary to a collaborative study of the immunogenetics of inflammatory bowel disease we have examined critically the evidence that genetic factors contribute to disease susceptibility.

A fundamental issue is the relation between Crohn's disease and ulcerative colitis. Although the classification of these conditions together is convenient, its scientific basis is not proved. We suggest that inflammatory bowel disease represents a number of distinct conditions of multifactorial aetiology. Some, but not all, predisposing genetic and environmental factors may be shared between subtypes of disease. Present knowledge of the extent of disease heterogeneity, however, is incomplete: this has contributed to the difficulty in defining the genetic basis of inflammatory bowel disease.

Evidence for genetic predisposition
A number of lines of study may be cited to show that genetic predisposition is important in susceptibility to inflammatory bowel disease, especially Crohn's disease. These include studies of disease prevalence in ethnic groups and migrants, and of the familial occurrence of disease. Critical review of these data suggest genetic factors interact with environmental factors: the comparative importance of nature and nurture is unclear.

ETHNIC AND MIGRANT GROUPS
Considerable differences exist in the prevalence of inflammatory bowel disease between different racial and ethnic groups. Crohn's disease and ulcerative colitis are most common in white people: lower prevalence rates are reported for native Americans, black Americans, Latin Americans, Maoris, and Asians. Among white groups, the ethnic group with the highest incidence and prevalence of inflammatory bowel disease are the Ashkenazic Jews of western Europe, United States, and Cape Town. Disease prevalence has been studied in Jewish migrants to Israel, and Asian migrants to the United Kingdom. In Israel, the Ashkenazic Jews who have migrated from Europe continue to show an increased disease prevalence, compared with the Sephardic Jews from north Africa and the Middle East. The overall prevalence of inflammatory bowel disease in Tel Aviv (in all ethnic groups), however, is much lower than that reported in western European or North American cities.

In the United Kingdom, the incidence and prevalence of inflammatory bowel disease has been studied in Asian migrants, and their offspring, who have settled in industrial cities (London, Bradford, Leicester). These studies suggest that these immigrants have an increased susceptibility to inflammatory bowel disease, compared with Asians resident in Asia. In particular, ulcerative colitis seems more prevalent in South Asians in Leicester, than in the native white population.

Environmental together with genetic factors are implicated by these studies: it does seem that both contribute to the development of all phenotypes of inflammatory bowel disease.

FAMILIAL INFLAMMATORY BOWEL DISEASE
Many studies, using differing methodologies, have shown an increased prevalence of inflammatory bowel disease among relatives of patients with Crohn's disease and ulcerative colitis. First degree relatives are at the greatest risk, particularly siblings, but more distant relatives also display an increased prevalence of disease. Both Crohn's disease and ulcerative colitis may occur in the same family.

The prevalence of a positive family history is uncertain, as the estimates vary considerably from study to study. Many of the data implying high disease prevalence in first degree relatives may be criticised for using selected patient groups.

Farmer et al4 have reported the highest prevalence of affected relatives, in a study based at the Cleveland Clinic. Of 522 patients with Crohn's disease, 187 (35.2%) were shown to have an affected family member (87 (16.7%) first degree). Of 316 ulcerative colitis patients 29-4% had a positive family history (50 (15-8%) first degree). These data include, however, only patients with onset of disease before 21 years of age. If either ulcerative colitis or Crohn's disease represents a number of genetically different conditions, then disease with an early age of onset may represent a subgroup with a particularly strong genetic influence.

A number of reports showing a high preval-
Genetics of inflammatory bowel disease

The TWIN inflammatory bowel disease. All twins with ulcerative colitis, which is important in the pathogenesis of chronic inflammatory bowel disease, is brought forward by twin pairs (dizygotic discordant). Disease, survey the Eighty Yang et al9 of familial bowel disease, were 4.5% and 5.2%, respectively. Age was used for relatives with Crohn's disease, and ulcerative colitis, 1-6% and 2-9%, respectively. This bias has been recognised; recently Yang et al10 have attempted to define the precise risks of inflammatory bowel disease in relatives of Jewish and non-Jewish patients, cared for in southern California. Five hundred and twenty-seven patients (291 Jews; 236 non-Jewish) were questioned. Age specific incidence data were used to estimate lifetime risks. In the first degree relatives of non-Jewish probands, the lifetime risks for inflammatory bowel disease were 5-2 and 1-6%, when probands had Crohn's disease and ulcerative colitis respectively. These values were consistently lower than the corresponding risks for relatives of Jewish patients (7-8 and 4-5% respectively).

McConnell has studied the family histories of 336 patients attending Broadgreen Hospital, in Liverpool.3 Thirty one (18-8%) of 165 patients with Crohn's disease had an affected first degree relative: in 27 cases, a sibling was affected. Of patients with ulcerative colitis, 20 of 171 had a positive family history (14-6%); in 16 cases, a sibling was affected. It was more probable that relatives of Crohn's disease patients would have Crohn's disease than ulcerative colitis, but both diseases were present in many families.

We have recently reviewed the family histories of patients with Crohn's disease and ulcerative colitis attending the gastroenterology clinic in Oxford. Case notes were reviewed. Of 317 patients with Crohn's disease, 41 (13%) have a positive family history for inflammatory bowel disease. In 21 cases, the affected relative is a sibling (6-6%). Of 825 patients with ulcerative colitis, 90 (12%) have a positive family history. In 31 (3-75%), a sibling is affected; in 24 (2-9%) parents are affected.

Environmental factors are common to first degree relatives sharing the same home. Van Kruiningen et al propose that shared environmental stimulus accounts for familial Crohn's disease, with particular reference to two large French families, with many affected relatives.11 Disease is not common, however, in spouses of patients with inflammatory bowel disease, whereas blood relatives separated by geography and generations retain an increased prevalence of Crohn's disease and ulcerative colitis.

TWIN STUDIES

The argument that genetic factors are of importance in the pathogenesis of inflammatory bowel disease, particularly Crohn's disease, is reinforced by studies of twins. Most pertinent is the survey of the Swedish registry of twin births.32 Eighty twin pairs were identified in which at least one twin was known to have inflammatory bowel disease. All twins were brought up in the same environment. Eight of 18 monozygotic twin pairs were concordant for Crohn's disease, but only one of 26 dizygotic pairs. In the twin pairs with one proband having ulcerative colitis, one of 16 monozygotic twins were concordant for the disease, but all other 20 twin pairs (dizygotic or unknown zygosity) were discordant. No 'mixed disease' twin pair, either monozygotic or dizygotic, was reported in this study.

The concordance of monozygotic twins for Crohn's disease is strong evidence for genetic susceptibility in this form of inflammatory bowel disease. The calculated heritability for Crohn's disease in this twin study was greater than values reported for diabetes, duodenal ulcer, schizophrenia, hypertension or bronchial asthma. Furthermore, the differences in concordance rates between Crohn's disease and ulcerative colitis, and the absence in published works of a mixed disease twin pair suggest that these phenotypes of inflammatory bowel disease have a distinct genetic basis.

ANKYLOSING SPONDYLITIS AND PRIMARY SCLEROSING CHOLANGITIS

Ankylosing spondylitis,22 and primary sclerosing cholangitis23 are recognised to have strong associations with distinct HLA specificities (HLA-B27 and B8 DR3 respectively), although the pathogenesis of both diseases awaits clarification. Both ankylosing spondylitis and primary sclerosing cholangitis occur more frequently than predicted in patients with inflammatory bowel disease and their first degree relatives. The mechanism underlying the association is uncertain. In the absence of ankylosing spondylitis, or sclerosing cholangitis, HLA-B27 and HLA-B8 DR3 are not associated with inflammatory bowel disease.

Pathophysiology and genetic susceptibility

If genetic susceptibility is important in the pathogenesis of chronic inflammatory bowel disease, what physiological process is affected? This problem remains unsolved, although recent investigations have pointed at aspects of non-specific defence mechanisms as well as the regulation of the antigen specific immune response.

NON-SPECIFIC DEFENCE MECHANISMS

Hollander et al showed increased intestinal permeability in 11 patients with Crohn's disease, and their 32 healthy first degree relatives, using the marker polyethylene glycol 400.30 These workers propose that this abnormality permits increased antigen uptake from the gut lumen, providing the stimulus for chronic inflammation. Although this is an area of considerable controversy,31 these initial findings have received recent support.32

The complement system, which may participate in the non-specific and specific immune response, has also been examined in Crohn's disease. Earlier work33 suggesting an association between F allotype of C3 and small bowel Crohn's disease (but not ulcerative colitis) has been supplanted by the finding of an enhanced production of complement components by the small intestine in healthy first degree relatives of probands with Crohn's disease.34 Compared with healthy unrelated controls, relatives had a 40% increase in jejunal fluid concentration of C4. Concentrations of C3 and
factor B were similar in healthy relatives and controls. These results, if reproducible, would imply that a primary (quantitative or qualitative) abnormality of complement secretion may contribute to the chronic inflammation of Crohn's disease.

The mucin glycoproteins of the small and large intestinal mucosa are thought to play an important part in the normal intestinal function, and in non-specific defence mechanisms. Using monoclonal antibodies directed against human colonic determinants, and detailed structural analysis, Podolsky et al 10 have classified human mucin glycoprotein species I to VI. These may reflect the products of functionally distinct sub-populations of colon goblet cells. Mucin isolated from colonic tissue from patients with Crohn's disease contained subclasses in similar proportions to those seen in normal controls. Mucin glycoproteins from patients with ulcerative colitis, however, were selectively deficient in subclass IV. 11 Recently evidence has emerged to suggest that this deficiency may be a primary mechanism, increasing host susceptibility to disease. 12 In monozygotic twin pairs, in which one sibling had ulcerative colitis, both siblings had reduced mucin subclass IV.

Molecular genotyping of mucin and complement genes is now available; these loci provide candidate genes for association studies, and linkage analysis (see below).

Genetics and the immune response

The activated cells of the immune system and their products are central to the pathogenesis of inflammatory bowel disease. Whether the chronic inflammatory response is an appropriate consequence of stimulation by exogenous antigen or a primary (genetically determined) defect of immunoregulation is uncertain. The second possibility has stimulated research into all aspects of immune function in patients with inflammatory bowel disease.

The key interaction in the generation of the specific immune response is that between the T cell receptor for antigen and the antigenic peptide, presented in the context of HLA molecules on the cell surface of antigen presenting cells. Processed antigen, presented in the context of class I HLA antigens will stimulate clonal proliferation and differentiation of specific cytotoxic lymphocytes. In contrast, 'helper' T cells will recognise antigen only in the context of class II antigens.

THE HLA SYSTEM

The structure of the glycoprotein molecules that make up class I and class II antigens is genetically encoded in the major histocompatibility complex on the short arm of chromosome six. Molecular analysis has established that these genes display a high degree of allelic polymorphism, not shown by serological methods alone.

A large number of investigators have sought association between ulcerative colitis and Crohn's disease and allelic variations in the genes of the major histocompatibility complex. In retrospect, methodological discrepancies are evident, particularly the reliance on serological typing of histocompatibility antigens rather than molecular genotyping. Early studies were hampered by incomplete knowledge of the HLA system, inadequate controls for ethnic differences, and small numbers of patients studied.

With the exception of patients also having ankylosing spondylitis, or primary sclerosing cholangitis noted above, the results have been largely inconclusive. No consistent HLA association has been reported using serological techniques in Crohn's disease, although meta analysis of published results of 730 pooled white patients showed that HLA-A2 allele carries a comparative risk of 1:25, whereas HLA-A11 has a significant negative association. 25

For ulcerative colitis, seven studies using serological techniques 26-32 have shown an increase in frequency of HLA-DR2 allele, although only two reached statistical significance. 26,32 The most recent studies of class II association have been designed successfully to overcome earlier difficulties, and provide encouragement. Toyoda et al 24 have conducted an association study using carefully selected ethnically matched cases and controls, and a control group of monozygotic molecular and serological techniques. The authors show separate genetic susceptibilities for Crohn's disease and ulcerative colitis. The haplotype HLA-DR1 DQ 6 was present in 27% of patients with Crohn's disease, compared with 13% controls (p=0.021). In contrast, HLA-DR2 was found to be present in 41% of patients with ulcerative colitis, compared with 21% controls, a statistically significant difference. Negative associations with DR4 and DR 6 were found in ulcerative colitis.

Although the early studies of HLA association with inflammatory bowel disease are generally regarded as inconclusive, it is relevant, in the light of this recent study, to re-evaluate them, and consider if particular phenotypes of disease may have defined genotypic association. The data from McConnell et al 26 and Asakura et al 33 both show that the association of HLA-DR2 is most pronounced with cases with total colonic involvement. Hence, specific genetic associations with phenotypically distinct subclasses of disease (for example, total colitis, proctitis) may be hidden, if ulcerative colitis is treated as genetically homogenous. Even if clinically apparent phenotypes of disease do not have a distinct genetic basis, subclasses may be defined by subclinical markers. Considerable recent interest concerns the use of the anti-neutrophil cytoplasmic antibody (ANCA) in defining subclasses of ulcerative colitis that are ANCA positive and ANCA negative. At present, however, results from North American 34 and western European 35 centres show marked discrepancies.

NON-HLA GENES

There are a number of non-HLA genes participating in the immune response. The genes that encode immunoglobulins and cytokines display allelic polymorphisms, which may be important in immunogenetic susceptibility to immune mediated disease. Of particular interest are the
Genes encoding the chains of the antigen receptor on the cell surface of T lymphocytes. This comprises two variable clonally distributed glycoproteins (either an αβ or γδ pair). The α, β, γ, δ chains are members of the immunoglobulin supergene family. Each protein chain has identifiable variable (v), joining (j), and constant (c) regions (β and δ chains also incorporate a diversity region (d) between v and j). The genes encoding these chains consist of a variable number of distinct segments representing the regions of the protein. Recombination of these segments permits for considerable diversity of T cell receptor structure (fully reviewed in reference 36). These genes are highly polymorphic and alleles may be defined by restriction fragment length polymorphisms (RFLP), or tandem repeat sequences. Preliminary studies have sought associations between alleles defined by RFLP and inflammatory bowel disease. Markers in the constant regions of both α and β chains have been investigated. These studies are prone to be limited by the failure to consider the lack of linkage disequilibrium across the region.

The repertoire of antigen receptors expressed on peripheral lymphocytes is determined not only by the diversity permitted by the T cell receptor genes, but also depends on subsequent events, notably intrathymic selection and deletion. Recent studies have indirectly implicated class II antigens and non-inherited maternal HLA antigens in determining expressed T cell repertoire in diabetes mellitus and rheumatoid arthritis respectively. Posnett et al. have shown oligoclonal T cell populations in mesenteric lymph nodes from resected specimens from Crohn's disease patients. This oligoclonality of Vβ receptor expression was not linked to a known polymorphism of the Vβ gene. Other workers are also exploring the expressed T cell repertoire in Crohn's disease.

**LINKAGE ANALYSIS**

Advances in molecular biology have increased our ability to define accurately disease associations with alleles at a given genetic locus. Although association studies may help in our understanding of the heterogeneity of inflammatory bowel disease and of the relation between ulcerative colitis and Crohn's disease, these population based studies are of limited use in precisely locating susceptibility genes. False positive associations may arise because of the use of an inappropriate control population; moreover, the gene identified in an association study may simply be in linkage disequilibrium with a true susceptibility gene.

To identify a true susceptibility gene, it is necessary to show linkage between disease and a candidate gene: inheritance of the gene must be shown to be significantly linked to disease inheritance.

In addition to the methodological difficulties inherent in association studies, linkage analysis provides fresh problems. To gain adequate statistical power, access is required to a large number of informative families, in which there is a strong history of inflammatory bowel disease. Furthermore, parametric methods of linkage analysis are most relevant to diseases in which the mode of inheritance, and extent of disease heterogeneity are clearly defined. In the light of these problems, detailed linkage analysis has not been carried out in inflammatory bowel disease, although a few small studies have been published (reviewed briefly in reference 24). The results of these early investigations do not permit accurate assessment of the overall genetic contribution to disease susceptibility.

At present independent investigators in the United States and Great Britain (in London and Oxford) are concentrating on the establishment of a repository of genetic material from inflammatory bowel disease families. Once this resource is available, formal linkage analysis using parametric and non-parametric methods (affected sibling pair) will be possible. As discussed, a number of candidate genes have presented themselves, in particular HLA class II, T cell receptors, complement, and mucin genes. It is hoped to evaluate accurately the contribution of these loci to overall disease susceptibility (as is already possible in a number of 'oligogenic diseases').

**Genetic heterogeneity and inflammatory bowel disease**

In conclusion, the classification of Crohn's disease together with ulcerative colitis as inflammatory bowel disease is convenient, but has a limited proved aetiological basis. Distinct pathological and physiological differences have been shown between these conditions, and genetic marker studies now also suggest distinct susceptibility. Any model proposed to explain disease inheritance, however, must also explain the fact that both diseases may coexist in one pedigree.

Crohn's disease and ulcerative colitis probably do not have a simple Mendelian mode of inheritance. The empiric risk to siblings and offspring shown in the familial studies are not in keeping with autosomal dominant or recessive inheritance. Complex segregation analysis, carried out on German and Scandinavian populations has suggested that a simple autosomal gene may be operating only in small subgroups of inflammatory bowel disease patients. Most recently, Orholm et al. assessed patterns of disease segregation in 637 Danish patients and their relatives. Analysis suggested an important dominant gene with a penetrance of 0·20-0·26 in 9-13% of adult patients with ulcerative colitis. For Crohn's disease, an autosomal recessive model was relevant in 7% of families.

An alternative model is the 'multiple loci - single disease model', applied to explain the familial association of Crohn's disease and ulcerative colitis. This model suggests that one genotype, with perhaps 10 to 15 genes, confers susceptibility to all forms of inflammatory bowel disease. If a person has an incomplete genotype, ulcerative colitis is more probable; if the genotype is complete Crohn's disease occurs. This model would account for the prevalence of ulcerative colitis in families of probands with Crohn's disease, and the stronger family history...

17 May GR, Sutherland LR, Meddings JB. Is small intestinal permeability really increased in relatives of patients with Crohn’s disease? Gastroenterology 1993; 104: 1627-32.


46 Hollander D. Permeability in Crohn’s disease: altered barrier