

Commentary

Linking genotype with phenotype in inflammatory bowel disease – Will we ever have reagent standard patients?

Subrata Ghosh*

Gastrointestinal Unit, Department of Medical Sciences, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, Scotland, UK

Received 8 January 2001

Keywords: Genotype, phenotype, IBD

Genetic factors predispose to a number of gastrointestinal diseases. The spectrum of genetic determination may vary from classical inherited diseases such as haemochromatosis where the genetic mutation is now fairly well defined to complex polygenic diseases, such as inflammatory bowel disease and coeliac disease. In inflammatory bowel disease, genetic predisposition confers disease susceptibility, but epigenetic factors (lifestyle or environmental triggers) are required to manifest the disease phenotype. The relatively rapid rise in incidence of inflammatory bowel disease especially Crohn's disease in some populations, such as children in Scotland [1], suggest that environmental triggers are likely to be very important, as the genotype of a cohort is unlikely to change so rapidly. The Scottish data confirms rapid temporal changes implicating environmental factors in the aetiology of both Crohn's disease and ulcerative colitis, and the relationship between the temporal changes of Crohn's disease and ulcerative colitis suggest certain shared risk factors. Preliminary prospective data for paediatric patients in the whole of UK show the same trend [2]. As the in-

cidence of Crohn's disease has now overtaken that of ulcerative colitis, additional risk factors for the expression of Crohn's disease are likely to exist as suggested previously by other authors [3]. The genetic basis of inflammatory bowel disease is however undisputed, as evidenced by familial clustering and increased concordance of the phenotype in monozygotic twins.

Systematic genome wide linkage analysis have defined genomic susceptibility regions in both Crohn's disease and ulcerative colitis [4,5]. Identification of specific risk mutations in genes within the relatively large linkage intervals is however, a much more complicated task. Specific genetic mutations of the genes encoding the cytokines involved in modulation of inflammation, such as IL-1 β , IL-1ra, TNF- α or IL-10 have naturally attracted intense attention, but with conflicting and varying results. Such relatively simple hypothesis driven research identifies key protein molecules involved in the disease pathogenesis, establishes genetic polymorphism of the gene(s) coding for the proteins, and then investigates association of a specific polymorphism with disease expression. However, there is considerable redundancy in the inflammatory pathways, and such hypothesis driven research limits investigation to known genes and pathways. Novel genes and regulatory mechanisms cannot be elucidated by such hypothesis driven research.

It is also relatively straightforward to identify genes in a monogenetic disorder in which inheritance follows Mendelian rules. Identification of the HFE gene in haemochromatosis is an example, though defining the exact function of the HFE gene still continues. Analysis of a few large families may lead to localisation of the disease gene, e.g. FAP in familial adenomatous polyposis. In complex polygenic disorders, with numerous phenotypic subgroups, such as inflammatory bowel disease, multiple disease genes are involved, and

*Correspondence to: Dr Subrata Ghosh, MD, FRCP(E), Consultant Gastroenterologist, Western General Hospital, Edinburgh EH4 2XU, Scotland, UK. Tel.: 0131 537 1755; Fax: 0131 537 1007; E-mail: sg@srv0.med.ed.ac.uk.

linkage analysis using a population based approach is necessary.

In inflammatory bowel disease, a series of linkage analysis have been performed. Hugot et al. [4] described a region of chromosome 16 which has been confirmed by other workers [6–10]. This is currently labelled IBD1. Satsangi et al. [5] found a region on chromosome 12 (IBD2) which was also reported by other workers [11,12]. However, in a recent study which used families from all over Europe, linkage on chromosome 12 could not be replicated in Crohn's disease [13]. It is likely that the putative gene in this region has only a small influence on the risk of developing inflammatory bowel disease. Schreiber and colleagues [14] have contributed greatly to our knowledge in this area and described IBD3 on chromosome 6, confirmed by Rioux et al. [15]. It is more likely than not that more such regions will be described. The discordance between studies may be largely accounted for by the numbers of the families studied, the magnitude of impact of the genes in the region on overall disease susceptibility and population specific differences.

Gender differences are of interest in immune mediated diseases, though the influence of gender on the incidence of either ulcerative colitis or Crohn's disease is small [16,17]. As reported in this issue, the distribution of cLMW-PTP genotypes in females with Crohn's disease was significantly different from that of controls [18]. The distribution of cLMW-PTP genotypes in males with ulcerative colitis was again significantly different from that of controls [18]. This is a novel observation, which requires further corroboration in inflammatory bowel disease patients from other geographical areas. Gender differences in disease rates may be helpful in giving clues as to aetiology. Hormonal, occupational or lifestyle factors may account for such gender differences.

This study is a report of a chance observation, not one based on an *a priori* hypothesis. Small gender differences have been noted in epidemiological studies of inflammatory bowel disease incidence and prevalence. These gender differences are certainly not as striking as in other classical autoimmune diseases. False positive results are not uncommon in case control type studies on specific candidate genes. This is especially true if the number of inflammatory bowel disease patients is relatively small and the confidence intervals for the relative risks are relatively large, as in this study. Notwithstanding these reservations, when combined epidemiologic data in inflammatory bowel disease is scrutinised, women are slightly more likely to develop

Crohn's disease and in general they are at 20–30% greater risk than men. Conversely, men are slightly more likely to develop ulcerative colitis [19]. This mirrors the cLMW-PTP genotypes reported as described above. The various members of the Protein Tyrosine Phosphatase (PTP) family are ubiquitously distributed and play a key role in cell signalling. It is not difficult to envisage members of the PTP family as being key regulators of inflammation in the intestine.

Relating genetic factors to broad phenotypes such as gender is useful but not as complex as relating to more relevant inflammatory bowel disease phenotypic subgroups based either on anatomical involvement or on associated extra-intestinal manifestations. We do not know whether the pathogenesis of ulcerative colitis and Crohn's disease is distinct, despite shared genetic predisposition. Some of the phenotypic features of inflammatory bowel disease lend themselves readily to 'reagent standard' grouping such as primary sclerosing cholangitis or inflammatory bowel disease associated arthropathies. However, others, such as the pharmacogenetics of response to corticosteroid therapy may be more complex – steroid resistance may be a phenotypic feature quite independent of inflammatory bowel disease. Recently, the Multidrug Resistance Gene (MDR) has been linked with failure of medical therapy in inflammatory bowel disease patients, with high expression of p-glycoprotein 170 both in gut mucosa and peripheral blood lymphocytes related to such therapeutic failure [20]. Anatomical extent of disease and complications often evolve with time in both ulcerative colitis and Crohn's disease.

Candidate gene studies are closer to phenotype as the choice of the gene is often based on phenotypic features and pathophysiological hypothesis. Polymorphisms in the population of the candidate genes lead to case-control type studies, but such studies are extremely susceptible to false positive findings because of population stratification. Cytokine, HLA and mucin genes provided a rich repertoire of candidate genes. IL-1 β polymorphisms [21] have been related to the severity of inflammatory bowel disease phenotype and polymorphisms controlling high IL-10 synthesis have been associated with inflammatory bowel disease [22]. Some of these findings contradicted earlier work, showing stringent study design is a very important factor in influencing conclusions. A case in point is the variable number of tandem repeats (VNTR) in the IL-1ra gene with conflicting results in different populations of ulcerative colitis patients [23–25]. In Crohn's disease, the results again have been very inconsistent re-

garding tumour necrosis factor (TNF) promoter polymorphisms and the flanking microsatellites [26–28]. A meta-analysis of HLA class I and II genes suggested the presence of a susceptibility gene on chromosome 6p [29,30]. Chromosome 6 is clearly well recognised as a very important chromosome for immune regulation. A wide range of other candidate genes have been investigated including mucin genes MUC2 and MUC3, heat shock protein HSP70-2, the IL-4 receptor NRAMPI, the interferon- γ and interferon- γ receptor genes and kinin B1 receptor.

Most clinicians would try to subgroup inflammatory bowel disease into two distinct types, i.e. ulcerative colitis and Crohn's disease. The therapeutic utility of this distinction is not huge as most of the therapeutic strategies for ulcerative colitis and Crohn's disease are broadly similar. However, the pharmacogenetic implication of phenotypic subgroups can be much more profound, eg. approximately 65% of severely ill Crohn's disease patients showed a dramatic response to infliximab (chimeric monoclonal antibody to TNF- α), whereas 35% with similar clinical characteristics and disease activity did not respond. This raises the hypothesis that there might be (at least) two forms of Crohn's disease, with different inflammatory processes at a molecular level. Apart from anatomical extent and extraintestinal features defining subgroups, serological subgroups based on anti-*Saccharomyces cerevisiae* antibody (ASCA) and anti-neutrophil cytoplasmic antibody with perinuclear staining pattern (pANCA) have been proposed. ASCA has been demonstrated in 50–60% of patients with Crohn's disease, but in only 10–15% of patients with ulcerative colitis. pANCA is found in 70% of patients diagnosed to have ulcerative colitis and only in 10–15% of Crohn's disease patients. These tests are considered potentially useful markers for distinguishing ulcerative colitis from Crohn's disease, especially in paediatric practice [31,32], but the specificity is too low for routine clinical use. Smoking and appendicectomy not only provide clues to pathogenesis, but might also serve to separate phenotypically distinct groups.

Certain subgroups of inflammatory bowel disease however do have striking gender differences in prevalence. Primary sclerosing cholangitis associated with ulcerative colitis (and, less often, Crohn's disease) typically occurs in males [33]. Determination of cLMW-PTP may be interesting in this 'gender enriched' subgroup. In contrast, the striking male preponderance of idiopathic ankylosing spondylitis is abolished in inflammatory bowel disease where the sex ratio is close to unity.

It is not known however, whether the phenotypic features are integral to inflammatory bowel disease or simply represent genotypic features unrelated to inflammatory bowel disease interacting with inflammatory bowel disease expression. It is known that after the first course of steroids in Crohn's disease for one month, 48% of patients obtained complete clinical response, 32% partial response and 20% were resistant and therefore had no response [34]. After steroids were tapered off, 44% of the total group maintained their response [34]. It is not known whether steroid unresponsiveness in a patient is a phenotypic subgroup of inflammatory bowel disease, or whether similar steroid unresponsiveness would be manifest in the patient for other diseases such as asthma or eczema. More than one gene almost certainly mediates steroid resistance.

The principal difficulty is in deciding whether distinct phenotypic features represent stable, homogeneous subgroups. Patients initially responsive to steroids may later behave as resistant. Stricturing Crohn's disease may also have associated inflammation and/or fistulae. Left sided ulcerative colitis may extend more proximally in a substantial proportion of patients. Extraintestinal manifestations occasionally might present later in the course of the disease.

We may never be able to have 'phenotypically pure' patients in a complex evolving disease such as inflammatory bowel disease, and therefore stable phenotypic features need to be identified which may form the basis of subgroups. Such features may be based on extraintestinal manifestations such as sclerosing cholangitis, arthropathy etc or newer serological markers such as ASCA or pANCA. Response to the newer biological therapies will give rise to other attractive subgroups for genetic associations. Such essentially case-control type studies are likely to throw up false positive associations due to population heterogeneity, and lack of power. High throughput 'hypothesis free' technologies have the greatest likelihood of success in complex disorders like IBD, and may eventually define genotypically different subgroups showing clear disease heterogeneity, such as the situation with 'essential' hypertension. Genomic expression analysis technology gives us a powerful technology for assessment of gene expression of potentially all human genes in a parallel and robotised automated fashion. Hybridisation based expression analysis techniques are commonly used, but an alternative representational cloning approach such as serial analysis of gene expression (SAGE) is also likely to be applied to inflammatory bowel disease genomics in future. Any attempt to phenotypically define

complicated subgroups in a complex evolving disease such as inflammatory bowel disease and construct large databases may be potentially wasteful effort currently, apart from a few exceptions mentioned above. Meanwhile, reports of genetic associations in inflammatory bowel disease, such as the cLMW-PTP reported in this issue may continue to provide clues about genes with a modest influence on disease expression and highlight phenotypic differences (such as gender) largely ignored in the past in inflammatory bowel disease.

References

- [1] E. Armitage, H. Drummond, S. Ghosh and A. Ferguson, Incidence of juvenile-onset Crohn's disease in Scotland, *Lancet* **353** (1999), 1496–1497.
- [2] B. Sandhu and A. Sawczenko, Inflammatory bowel disease (IBD) in under 20-year-olds. Royal College of Paediatrics and Child Health/ British Paediatric Surveillance Unit, 14th Annual Report, 1999–2000, pp. 25–27.
- [3] F. Delco and A. Sonnenberg, Commonalities in the time trend of Crohn's disease and ulcerative colitis, *Am J Gastroenterol* **94** (1999), 2171–2176.
- [4] J.P. Hugot, P. Laurentpuig and C. Gower-Rousseau et al., Mapping of a susceptibility locus for Crohn's disease on chromosome 16, *Nature* **379** (1996), 821–823.
- [5] J. Satsangi, M. Parkes and E. Louis et al., Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosome 3, 7 and 12, *Nature Genet* **14** (1996), 199–202.
- [6] J.D. Ohmen, H.Y. Yang and K.K. Yamamoto et al., Susceptibility locus for inflammatory bowel disease on chromosome 16 has a role in Crohn's disease, but not in ulcerative colitis, *Hum Mol Genet* **5** (1996), 1679–1683.
- [7] J.H. Cho, Y. Fu and B.S. Kirshner et al., Confirmation of a susceptibility locus, IBD1, for Crohn's disease on chromosome 16, *Inflamm Bowel Dis* **3** (1997), 186–190.
- [8] J.A. Cavanaugh, D.F. Callen and S.R. Wilson et al., Analysis of Australian Crohn's disease pedigrees refines the localization for susceptibility to inflammatory bowel disease on chromosome 16, *Ann Human Genet* **62** (1998), 291–298.
- [9] V. Annesse, A. Latiano and P. Bovio et al., Genetic analysis in Italian families with inflammatory bowel disease supports linkage to the IBD1 locus—a GISC study, *Eur J Hum Genet* **7** (1999), 567–573.
- [10] S.R. Brant, Y. Fu and C.T. Fields et al., American families with Crohn's disease have strong evidence for linkage to chromosome 16 but not chromosome 12, *Gastroenterology* **115** (1998), 1056–1061.
- [11] H. Yang, J.D. Ohmen and R. Ma et al., Linkage and association between Crohn's disease and a putative locus on chromosome 12, *Gastroenterology* **114** (1998), A902.
- [12] R.H. Duerr, M. Barnada and L. Zhang et al., Linkage and association between inflammatory bowel disease and a locus on chromosome 12, *Am J Hum Genet* **63** (1998), 95–100.
- [13] S. Lesage, H. Zouali and J.-F. Colombel et al., Genetic analysis of chromosome 12 loci in Crohn's disease, *Gut* **47** (2000), 787–791.
- [14] J. Hampe, S.H. Shaw and R. Saiz et al., Linkage of inflammatory bowel disease to human chromosome 6p, *Am J Hum Genet* **65** (1999), 1647–1655.
- [15] J.D. Rioux, M.S. Silverberg and M.J. Daly et al., Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci, *Am J Hum Genet* **66** (2000), 1863–1870.
- [16] H. Nyhlin and A. Danielsson, Incidence of Crohn's disease in a defined population in northern Sweden, 1974–1981, *Scand J Gastroenterol* **21** (1986), 1185–1192.
- [17] V. Binder, H. Both and P.K. Hansen et al., Incidence and prevalence of ulcerative colitis and Crohn's disease in the county of Copenhagen, 1962–1978, *Gastroenterology* **83** (1982), 563–568.
- [18] N. Bottini, F. Gloria-Bottini, N. Lucarini, P.G. Ronchetti and L. Fontana, Inflammatory bowel disease: Are there gender differences in the genetics of signal transduction? A preliminary study of cytosolic low molecular weight protein tyrosine phosphatase, *Disease Markers* **16** (2000), 163–166.
- [19] R.S. Sandler, Epidemiology of inflammatory bowel disease, in: *Inflammatory Bowel Disease -From Bench to Bedside*, S.R. Targan, S.F. Williams and Wilkins, eds, Baltimore, Maryland, USA, 1994, pp. 5–30.
- [20] R.J. Farrell, A. Murphy and A. Long et al., High multidrug resistance (p-glycoprotein 170) expression in inflammatory bowel disease patients who fail medical therapy, *Gastroenterology* **118** (2000), 279–288.
- [21] A. Nemetz, M.P. Nosti-Escanilla and T. Molnar et al., IL1B gene polymorphisms influence the course and severity of inflammatory bowel disease, *Immunogenetics* **49** (1999), 527–531.
- [22] A. Tagore, W.M. Gonsalkorale and V. Pravica et al., Interleukin-10 (IL-10) genotypes in inflammatory bowel disease, *Tissue Antigens* **54** (1999), 386–390.
- [23] J.C. Mansfield, H. Holden and J.K. Tarlow et al., Novel genetic association between ulcerative colitis and the anti-inflammatory cytokine interleukin-1 receptor antagonist, *Gastroenterology* **106** (1994), 637–642.
- [24] P.C. Stokkers, B.E. van Aken and N. Basoski et al., Five genetic markers in the interleukin 1 family in relation to inflammatory bowel disease, *Gut* **43** (1998), 33–39.
- [25] N.A. Tountas, V. Casini-Raggi and H. Yang et al., Functional and ethnic association of allele 2 of the interleukin-1 receptor antagonist gene in ulcerative colitis, *Gastroenterology* **117** (1999), 806–813.
- [26] S.E. Plevy, S.R. Targan and H. Yang et al., Tumour necrosis factor microsatellites define a Crohn's disease-associated haplotype on chromosome 6, *Gastroenterology* **110** (1996), 1053–1060.
- [27] G. Bouma, B. Xia and J.B. Crusius et al., Distribution of four polymorphisms in the tumour necrosis factor (TNF) genes in patients with inflammatory bowel disease (IBD), *Clin Exp Immunol* **103** (1996), 391–396.
- [28] K. Negoro, Y. Kinouchi and N. Hiwatashi et al., Crohn's disease is associated with novel polymorphisms in the 5' -flanking region of the tumour necrosis factor gene, *Gastroenterology* **117** (1999), 1062–1068.
- [29] P.C. Stokkers, P.H. Reitsma, G.N. Tytgat and S.J. van Deventer, HLA-DR and -DQ phenotypes in inflammatory bowel disease: a meta analysis, *Gut* **45** (1999), 395–401.
- [30] H. Yang, S.E. Plevy and K. Taylor et al., Linkage of Crohn's disease to the major histocompatibility complex region is detected by multiple non-parametric analyses, *Gut* **44** (1999), 519–526.
- [31] F. Ruemmele, S. Targan and G. Levy et al., Diagnostic accuracy of serological assays in pediatric inflammatory bowel disease, *Gastroenterology* **115** (1998), 822–829.

- [32] E. Hoffenberg, S. Fianza and A. Sauaia, Serologic testing for inflammatory bowel disease, *J Pediatr* **134** (1999), 447–452.
- [33] S. Ghosh, A. Shand and A. Ferguson, Ulcerative colitis, *BMJ* **320** (2000), 1119–1123.
- [34] P. Munkholm, E. Langholz, M. Davidsen and V. Binder, Frequency of glucocorticoid resistance and dependency in Crohn's disease, *Gut* **35** (1994), 360–362.