

Gene–Gene Interactions in Inflammatory Bowel Disease: Biological and Clinical Implications

Jean-Paul Achkar, MD, FACP^{1,2} and Claudio Fiocchi, MD^{1,2}

Abstract: The results of recently completed genome-wide association studies have advanced knowledge of inflammatory bowel disease (IBD) genetics, including the identification of over 30 genes or loci associated with Crohn's disease (CD). The possibility of interactions between genes, referred to by the term *epistasis*, needs to be carefully considered as both genetic and functional studies in IBD move forward. We review a paper in this issue of the Journal that reports evidence of epistasis in CD, and we discuss important issues that arise when trying to determine whether there is indeed interaction between genes and what their potential implications for disease pathogenesis and clinical outcome might be.

Am J Gastroenterol advance online publication, 12 May 2009;
doi:10.1038/ajg.2009.179

Since the discovery in 2001 of the first inflammatory bowel disease (IBD)-associated gene, the nucleotide-binding oligomerization domain 2 (*NOD2*) gene (1,2), our understanding of the genetics of Crohn's disease (CD) and ulcerative colitis (UC) has rapidly evolved. Most recently, progress in high-throughput genotyping technology and increased knowledge about the human genome through the International HapMap Project and the Human Genome Project have enabled the completion of genome-wide association studies for several complex diseases, including IBD. Such studies have led to the expansion of our knowledge of IBD genetics, leading to the identification of over 30 CD-associated genes or loci (3) and a growing number of UC-associated genes (4,5). The genes implicated in such studies, including *IL23R*, *IL12B*, *IL10*, *NOD2*, *ATG16L1*, and *IRGM*, have highlighted

the important roles of the immune system and of the intestinal tract's interaction with and response to the gut microflora in the pathogenesis of IBD. We have also learned that most of the identified genes individually have only modest effects on IBD susceptibility, suggesting that IBD is most likely the result of a combination of several genetic alterations as well as other factors such as environmental influences.

Given the association of over 30 genes or loci with CD and the finding that most of these regions only have modest effects on CD susceptibility, it is plausible that there are interactions between genes that further affect the risk and clinical course of CD. One possibility is that such interactions, often referred to as epistasis, are ubiquitous among common human diseases and even more important than the independent effects of any single susceptibility gene (6). A PubMed search using the words "epistasis" and "IBD" yields close to 30 publications, underscoring a substantial amount of interest and active investigation in gene–gene interactions in IBD. To further expand our understanding of IBD genetics, potential gene–gene interactions will need to be carefully delineated, especially if they lead to functional alterations or different phenotypic expressions of disease that might otherwise go unrecognized. In addition, the influence of environmental factors such as possible environment–gene interactions or modifications also need to be considered.

In the study by Torok *et al.* (7) in this issue of the Journal, the authors follow-up on an earlier report in which they found an association between a polymorphism in the Toll-like receptor 9 (*TLR9*) gene and CD (8). In addition, the authors tried to determine whether there are interactions between the *TLR9* variant and other CD-associated genes. Toll-like receptors are a family of pattern recognition receptors that play an important role in host defense by recognizing conserved molecular products of microbes and initiating innate immune responses (9). Some studies have shown possible interactions between *NOD2* and TLR proteins, including a recent study showing that functional synergy between

¹Department of Gastroenterology and Hepatology, Digestive Disease Institute, Cleveland Clinic, Cleveland, Ohio, USA; ²Department of Pathobiology, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA. **Correspondence:** Jean-Paul Achkar, MD, FACP, Department of Gastroenterology, Cleveland Clinic, 9500 Euclid Avenue, Desk A30, Cleveland, Ohio 44195, USA. E-mail: achkarj@ccf.org

Received 6 February 2009; accepted 17 February 2009

TLR9 and NOD2 is lost in CD patients who are homozygous for *NOD2* variants (10).

This study uses a case-control design involving German Caucasians and includes both the initial pilot population of 312 IBD subjects (174 CD, 138 UC) and 265 controls, as well as a replication cohort of 644 IBD subjects (432 CD, 212 UC) and 527 controls (7,8). The authors first found that, using the replication cohort, they were unable to confirm their earlier finding of an association of the *TLR9*-1237 T/C polymorphism with CD. However, on stratified analysis of the *NOD2* variant, a significant increase in the frequency of the *TLR9*-1237 T/C polymorphism was detected among CD patients with at least one *NOD2* variant compared with those with a wild-type *NOD2* genotype (odds ratio=1.74, confidence interval=1.19–2.54). Furthermore, there was a “gene dose related” effect with higher frequencies of *TLR9*-1237 T/C polymorphisms among those with two *NOD2* variants compared with those with one *NOD2* variant (odds ratio=2.77, confidence interval=1.42–4.68 compared with *NOD2*-negative subjects). They then assessed for further possible epistasis of this *TLR9* polymorphism with other previously described CD-associated variants using logistic regression. This analysis revealed interactions between the *TLR9* polymorphism and four *IL23R* variants and one *DLG5* variant, whereas weaker interactions were seen for *NOD2* and *ATG16L1*, but these did not maintain statistical significance after correction for multiple testing. Of note, the interactions with *IL23R* variants oscillated, with a lower frequency of the *TLR9* polymorphism for CD patients who had *IL23R* variants associated with increased CD risk and the converse for CD patients who had *IL23R* variants associated with a protective effect against the development of CD. The authors concluded that they have found the first evidence of significant genetic interactions between polymorphisms in *TLR9* and variants in *NOD2*, *IL23R*, and *DLG5*.

In reviewing the study by Torok *et al.*, it is important to first note that the *TLR9* polymorphism itself was not independently associated with CD in this replication study nor has it been previously found to be associated with CD in other IBD genetic studies. Although this does not exclude the possibility that *TLR9* plays a role in IBD pathogenesis, replication is a crucial component of genetic studies in order to exclude the risk of spurious associations. In addition, although the reported finding of epistasis between *TLR9* and three of the five other CD-associated genes assessed in this study may suggest a ubiquitous effect for *TLR9*, the finding of this high degree of epistasis is somewhat puzzling, especially considering the finding of independent interactions of the *TLR9* polymorphism with multiple genes related to different response pathways (both immune and microbial). Despite these limitations, this study does bring the concept of epistasis to the forefront and allows us to begin to think about how to approach dissecting our way through other gene-gene interactions.

Although this study begins to scratch the surface of the topic of gene-gene interactions in complex human genetic diseases

such as IBD, it also raises some important cautionary points and questions. The authors use the word *epistasis*, a term broadly defined as interaction between genes at different loci, to report the main findings of this study. Of importance, the literature often uses variable definitions of the term *epistasis*, some referring to different mathematical tests that show evidence of departure from independence of the effects of genetic loci on disease susceptibility, others referring to biological interaction of protein products derived from the genes being studied, and yet others referring to interactions that lead to phenotypic changes (11,12). In this study, the term epistasis is being used to describe statistical/mathematical evidence of interaction between two genes. Because this is a case-control association study with a dichotomous trait (disease vs. no disease), the authors appropriately use logistic regression to conduct such an analysis (11). The authors also provide a good theoretical rationale for a possible role of aberrant *TLR9* function in development of IBD and reference earlier studies that detected functional interactions in bacterial sensing by *NOD2* and *TLR9*. However, the authors did not find evidence of significant phenotypic expression related to this interaction, although this part of the analysis was limited by small numbers, as only about half the study participants had full clinical characterization available. Certainly follow-up studies to first confirm the presence of this interaction in different study groups will be needed. If a statistical interaction is confirmed, it will then be imperative to determine whether there are functional or phenotypic consequences associated with this epistatic finding. It is also important to consider that functional follow-up studies related to epistasis will, in general, be challenging, as *in vivo* and *in vitro* biological studies are usually best designed to assess the difference of alterations in single parameters.

In this regard, what are, if any, the biological implications of epistasis? In principle, the possible consequences could be fairly straightforward, such as pathogenic, phenotypic, clinical, and response to therapy outcomes that are different from those anticipated from the input of single genes. This seemingly logical assumption is hinted at in the work of Torok *et al.* (7) when they mention the possibility that epistasis of *TLR9* with other genes may induce a more severe disease phenotype, create alternative pathways of CD pathogenesis, differentially modulate CD susceptibility, or influence the response to therapy. Unfortunately, these theoretical expectations will probably be hard to test and prove, given the nature and intrinsic limitations of epistasis as objectively discussed by Cordell (11). In fact, the detection of epistasis suggests that we are dealing with a biologically interesting phenomenon, but at this point it can only be assumed that epistasis will reveal new findings in regard to mechanisms of disease, such as interactions among biologically relevant molecules, similar to cytokines and their receptors, for example, in the case of an inflammatory process such as IBD. Of importance, statistical analyses of interaction are limited to testing hypotheses regarding quantities but not biological responses, as there is no precise correspondence between mathematically and biologically derived models

of epistasis, and statistical interactions do not imply interaction at a pathophysiological level (11). Empirical data across a range of traits and species imply that most genetic variance is additive (13), and a statistical correlation between two susceptibility genes that individually increase the risk of CD intuitively compel us believe that the risk for this form of IBD is further increased or that the clinical phenotype will be more severe. This instinctive but superficial deduction is based on the independent action of each gene *per se*, but one cannot necessarily expect the same effect when the two genes interact *in vivo* in the presence of other modifying genes and innumerable other factors derived from both the endogenous and exogenous environment. Considering the continuous progress of genome-wide association studies in humans, and the increasing number of genes with small effects on disease phenotype (14), the number of statistically derived epistatic interactions is certain to increase, making the question of how to interpret their significance and how to test their potential biological value an enormously difficult task. Innumerable experiments should be carried out in humans to confirm that all genetic variant-determined pathways do lead to functional abnormalities, and a whole new set of knockout, knock-in, and transgenic animal models should be generated to carry out finer experimentation not feasible in humans. Clearly not enough resources are now at hand to perform all that is required to study the currently known and future epistatic interactions to be uncovered by the ever increasing number of IBD-associated genetic variations. The use of bioinformatics may alleviate some of these problems, but this approach is still in its infancy and is largely untested for complex diseases such as IBD; its use may provide answers that still need to be tested by cellular and molecular methods rather than mathematical or *in silico* systems.

CONFLICT OF INTEREST

Guarantor of the article: Jean-Paul Achkar, MD, FACG.

Specific author contributions: Achkar and Fiocchi contributed equally to the writing of this editorial.

Financial support: None.

Potential competing interests: None.

REFERENCES

- Hugot JP, Chamaillard M, Zouali H *et al*. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001;411:599-603.
- Ogura Y, Bonen DK, Inohara N *et al*. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 2001;411:603-6.
- Barrett JC, Hansoul S, Nicolae DL *et al*. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008;40:955-62.
- Franke A, Balschun T, Karlsen TH *et al*. Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. *Nat Genet* 2008;40:1319-23.
- Silverberg MS, Cho JH, Rioux JD *et al*. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nat Genet* 2009;41:216-20.
- Moore JH. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. *Hum Hered* 2003;56:73-82.
- Torok HP, Glas J, Endres I *et al*. Epistasis between Toll-like receptor (TLR)-9 polymorphisms and variants in *NOD2* and *IL23R* modulates susceptibility to Crohn's disease. *Am J Gastroenterol*; advance online publication, 12 May 2009 [e-pub ahead of print].
- Torok HP, Glas J, Tonenchi L *et al*. Crohn's disease is associated with a toll-like receptor-9 polymorphism. *Gastroenterology* 2004;127:365-6.
- Rakoff-Nahoum S, Paglino J, Eslami-Varzaneh F *et al*. Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* 2004;118:229-41.
- Van Heel DA, Ghosh S, Hunt KA *et al*. Synergy between TLR9 and NOD2 innate immune responses is lost in genetic Crohn's disease. *Gut* 2005;54:1553-7.
- Cordell HJ. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum Mol Genet* 2002;11:2463-8.
- Nagel RL. Epistasis and the genetics of human disease. *C.R. Biologies* 2005;328:606-15.
- Hill WG, Goddard ME, Visscher PM. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet* 2008;4:e1000008.
- Donnelly P. Progress and challenges in genome-wide association studies in humans. *Nature* 2008;456:728-31.