NIK SPENCER

GENETICS

Clues in the code

Gene exploration is providing unexpected insights into inflammatory bowel disease, and getting scientists closer to finding treatments that target the biological mechanisms.

BY SARAH DEWEERDT

Nicholas Volker first developed the symptoms of inflammatory bowel disease (IBD) shortly before his second birthday. Childhood cases of this gut disease are often severe, but even by this standard Nicholas's case was extreme. Over the next three years, his parents and doctors watched, helpless, as wounds opened up on his abdomen and leaked faeces. The doctors could do nothing as treatment after treatment failed, and Nicholas endured sepsis, excruciating pain and more than 100 surgical procedures.

In 2009, geneticist Howard Jacob, then at the Medical College of Wisconsin in Milwaukee, and his team sequenced Nicholas's genome in a last-ditch effort to find the cause of his symptoms and, they hoped, save his life. The team discovered that Nicholas had a mutation in the gene *XIAP*, which had previously been linked to immune deficiency. Doctors performed a cord-blood transplant, giving Nicholas a complement of stem cells that form immune cells with an intact *XIAP*, and his gut symptoms soon went into remission. His case is frequently touted as the first example of a person cured of any condition as a result of DNA sequencing.

Paediatric IBD is thought to be predominantly due to genetic factors. Researchers have identified around 50 genes that, when mutated, can each cause IBD symptoms in young children. Many of these, like *XIAP*, have also been linked to immune deficiency. In a large number of the youngest patients, IBD can be thought of as a rare, single-gene disorder.

In adults, the picture is much more complex — a clean fix is not an option. Over the past 15 years, research into the genetics of IBD has identified around 200 genomic hotspots that influence the risk of Crohn's disease or ulcerative colitis, the two major forms of IBD (see page S98). But so far, this information has had little clinical impact.

"These are flags waving towards different regions of the genome," says Sarah Ennis, who leads the Genomics Informatics Group at the University of Southampton, UK. "But they're not very useful for a consultant sitting with a patient in front of him or her." These genetic flags have provided clues to the underlying mechanisms of IBD, and revealed surprising subgroups within the IBD spectrum, connections with other conditions and possible drug targets. The challenge now, researchers such as Ennis say, is to apply this knowledge to improve the lives of patients.

SAME DIFFERENCE

This challenge is hard to meet, however. In adults, IBD results from a complex mixture of environmental risk factors, which are themselves not fully understood, and genetic factors that have varying impact (see page S100). A clear pattern that has emerged is that the strongest genetic risk factors tend to be specific to either Crohn's disease or ulcerative colitis. These differences provide clues to the mechanisms behind each disease and will hopefully

inform new treatment approaches.

The biggest player in Crohn's disease is a gene called *NOD2*, an important part of the innate immune system — the set of non-specific mechanisms that provide the body's first line of defence against infection. *NOD2*'s role was uncovered in 2001, when two independent groups of researchers traced the inheritance of gut problems in families affected by the disease[1,2]. "That was the first major insight that came from gene discovery, and that really spurred the next decade of immunological research into Crohn's disease," says Charlie Lees, a gastroenterologist at the University of Edinburgh, UK. The advance also sparked a surge of research into the genetics of IBD more broadly.

Subsequent research confirmed the link between *NOD2* and Crohn's disease, but found only a weak association with ulcerative colitis. Autophagy, the process by which cells process and degrade intracellular bacteria and cellular components, has also been implicated only in Crohn's disease. These findings suggest that this condition could be due to an abnormal immune response to microbiota, the community of bacteria that inhabits our guts.

For ulcerative colitis, the link seems to be with variations to genes involved in maintaining the integrity and function of the digestive-tract lining, suggesting that this disease might be caused by an inadequate barrier or 'leaky gut'. Variations within a set of genes from the human leukocyte antigen gene family (HLA class II) — which are involved in fine-tuning the immune system and helping it to recognize proteins made by foreign invaders — also make a major contribution to ulcerative-colitis risk, but are only weakly linked with Crohn's disease. This suggests that there might be an external trigger for ulcerative colitis, similar to the role that gluten has in coeliac disease, says Andre Franke, director of the Institute of Clinical Molecular Biology at the University of Kiel in Germany.

A DISEASE APART

Genetic investigations are also challenging the idea that IBD can be divided into two main diseases — Crohn's that can affect any part of the digestive tract and ulcerative colitis, which affects only the colon. In January, Lees and colleagues reported that certain genes determine where in the gut IBD causes inflammation, and that the genetic distinction provided a different way to classify IBD[3]. "What we found was that this disease location separated out IBD much more naturally into three main disease types, rather than just two," Lees says. "Small-bowel Crohn's disease was as distinct from colonic Crohn's disease as it was from ulcerative colitis."

Curiously, the three-disease classification is how doctors thought of IBD half a century ago, but rather than that system being superseded by an advance in biological understanding, it seems to have steadily drifted out of favour, Lees says. There's still substantial genetic overlap between the three groups, so the analysis is not immediately clinically applicable. But the findings do suggest that simply sorting people into either those with Crohn's disease or those with ulcerative colitis may not be enough to determine the best treatment for them.

Because so many genes linked with the immune system have been implicated in IBD, researchers are looking for genetic connections with other inflammatory and autoimmune diseases. About 70% of the genetic loci that contribute to IBD susceptibility are thought to be shared with other complex diseases such as ankylosing spondylitis.

Perhaps the most surprising link that genetic studies have thrown up is that between IBD and susceptibility to mycobacteria — a group that includes the pathogens that cause tuberculosis and leprosy. *NOD2* and roughly a dozen other genetic factors associated with IBD, particularly Crohn's disease, are also associated with susceptibility to tuberculosis and leprosy. There are two main ideas about the biological mechanisms behind this link.

"Genomic hotspots are flags waving towards different regions of the genome."

It is possible that certain genetic variants increase susceptibility to leprosy while also increasing the risk of Crohn's disease. "I believe that Crohn's disease is an immunodeficiency," Franke says. In people who are susceptible to leprosy, the immune system overlooks the bacterium and enables the infection to become established. Similarly, Franke explains, this deficient immune response may permit normal gut bacteria to penetrate the intestinal wall and trigger the inflammation seen in Crohn's disease.

Or, the link could run the other way around: some genetic variants might confer resistance to leprosy, but by coincidence increase the risk of Crohn's disease. Support for this hypothesis comes from the fact that some of the variants that increase IBD susceptibility are so widespread — present in 20–50% of people in certain populations — that many scientists think that they must have conferred some evolutionary advantage in the past.

FINE MAP

Many of these genetic insights into IBD have come from genome-wide association studies (GWAS), which scan the genomes of large numbers of people with the disease and compare certain genetic markers to those of controls. GWAS have been unusually successful in IBD, partly because of the uncommonly large and well-coordinated International Inflammatory Bowel Disease Genetics Consortium. "Around the globe we've got probably hundreds of clinicians contributing their own data sets and samples and patients into these studies," says Miles Parkes, a gastroenterologist at the University of Cambridge, UK, and a member of the consortium. Some of the largest analyses have involved more than 75,000 patients.

GWAS are a powerful approach — a practical way of finding markers for disease risk across the vast human genome — but they do have their downsides. For one, the gene chips (slides containing known sequences that are used to rapidly scan DNA for variations) are based on European genetics. But different ethnic groups are thought to have separate suites of genes that put them at risk. *NOD2* variants seem to have little involvement in Japanese and Korean people with Crohn's disease, for instance. And a study of more than 2,300 African Americans with IBD found 2 genetic markers linked to ulcerative colitis that have not been seen in people of European ancestry[4]. The European chips have complicated efforts to identify the genes that are most important in various ethnic groups. And GWAS for IBD have so far involved a disproportionate number of individuals of European ancestry.

Moreover, GWAS simply identify loci — positions in the genome — not the genes themselves, nor the precise mutations involved. On their own, GWAS don't directly illuminate the biological mechanisms behind the disease that might translate into clinical applications.

Researchers are making progress in identifying the specific genes that increase the risk of IBD. The pro-inflammatory molecule interleukin-23 (IL-23) and other proteins in its signalling pathway, for example, repeatedly pop up as players in both Crohn's disease and ulcerative colitis. Indeed, several pharmaceutical companies have drug candidates in clinical trials that target the IL-23 pathway. Many of the most promising IBD treatments in the pipeline involve biochemical pathways highlighted by genetics studies, says Parkes.

Still, a more comprehensive picture of the specific genes and variants involved is needed to move new treatments for IBD forward. And that picture may be about to develop. The International Inflammatory Bowel Disease Genetics Consortium has recently completed a three-year 'fine-mapping' effort that has pinpointed 45 specific genetic variants that contribute to the disease. The data suggest that some of these variants lie outside the coding and regulatory regions of genes — in parts of the genome that have no known function[5].

Working out how these variations contribute to disease will be a major challenge, but an exciting one. "There's clearly a huge amount of new biology waiting to be discovered," Parkes says. ■

Sarah DeWeerdt *is a freelance science writer based in Seattle, Washington.*

1. Hugot, J.-P. *et al. Nature* **411,** 599–603 (2001).
2. Ogura, Y. *et al. Nature* **411,** 603–606 (2001).
3. Cleynen, I. *et al. Lancet* **387,** 156–167 (2016).
4. Brant, S. R. *et al. Gastroenterology* http://dx.doi.org/10.1053/j.gastro.2016.09.032 (2016).
5. Huang, H. *et al.* Preprint at bioRxiv http://dx.doi.org/10.1101/028688 (2015).