The identification of a Crohn's disease gene

9/7/07

A team at the Giga-Research centre at the University of Liège has just isolated, in Chromosome 5, a genetic variant which renders some people susceptible to Crohn's disease. This chronic inflammation of the digestive system, mainly of the intestine and the colon, affects around 10,000 people in Belgium. The Liège discovery opens up a clearly signposted road down which the search for treatment must travel.

Comparing the genome of a person affected by Crohn's disease with that of someone who is unaffected is like placing side by side two books, each with a million pages which are identical save for a few letters, and looking for these infinitesimally minute differences without knowing on which page, or even in which chapter, these differences are hidden. Mission impossible? Not exactly! Researchers have just discovered a very precise genetic zone, based on Chromosome 5, which clearly plays a role in the appearance of this inflammatory disease of the digestive system. They published their results in the journal PLOS Genetics last April.

This discovery is first of all the result of nothing less than an army of ants going industriously about their work: building up a sufficiently large DNA bank of patients affected by Crohn's disease to allow for a statistically significant study. It has not been a question of integrating tens or hundreds of ill people into this database, but thousands! This is because Crohn's disease is a multi factored disease in which are mixed genetic predisposition and environmental factors. Thanks to his clinical work at the University Hospital Centre, in the department of Jacques Belaïche, who also made an active contribution to the work, and thanks also to the work of other gastroenterologists working in other hospitals, Dr. Edouard Louis patiently built up this databank: carrying out examinations of blood samples, intestinal biopsies, and painstaking encoding of medical case studies. Long years of work have been devoted to this job! In adding the databases of the universities of Louvain, Brussels and Gent to the Liège databank over 3000 people have been included in the complete database: 1700 patients and 1500 healthy people in order to construct a control group.

A meeting between two researchers

Very early on interested in genetics - he was trained in Professor Derek Jewell's laboratory at the University of Oxford in 1994 and 1995 - Edouard Louis crossed the path of another researcher at the university, Michel Georges, a specialist in animal genetics. The same generation, the same passion for research: they clicked and the two researchers decided to work together. The doctor brought to the table his databases and knowledge of Crohn's disease, the geneticist provided the very latest developments in genome research. This research had undergone a massive acceleration at the turn of the twenty-first century, which is why Michel Georges says that: 'From Mendel to genome sequencing is really the prehistoric age of genetics!' If, to enter History as such, humans discovered writing 5000 years ago, to enter the history of genetics they have learned how to decode the DNA alphabet and how to read the great book of life. Revealed in April 2003, the sequence is 99.99% precise. That is to say it contains a single error every 10,000 letters.

But it is one thing to decode the book of life, to know in basic terms the three million letters which compose a molecule of human DNA, and quite another to understand how the genome functions, and what the purpose of all these letters is. Since the completion of human genome sequencing thousands of researchers have hitched themselves up to this task. The first riddle: where are the genes, that is to say the genome's zones whose information is used by the cell to manufacture the proteins which are the organism's real workers? The genes only make up 2 to 3% of the whole of the genome. Before the results of human genome sequencing
were made known it was thought that the human genome contained around 100,000 genes. Today we know that the figure is much closer to 20,000. The second great mystery: what purpose is served by all the rest, the non encoding parts of the genome? At the start considered to be insignificant, the non encoding parts of the genome are today interesting scientists very much, notably the 'regulating' zones, which play a role in switching genetic expression on or off, or in any case a role in modulating it.

Another great leap forward in genetics since the sequencing of the human genome has been the discovery of SNPs, Single Nucleotide Polymorphisms. These are the places in the genome which differ very greatly from individual to individual, by a single letter or so. They represent more than 90% of all the genetic differences between individuals. Two humans picked out at random are 99.9% identical. The SNPs represent the great majority of the remaining 0.1% of differences. It is there that are housed a very large part of the differences between people; the tall, the short, the heavily set, the scrawny, etc. But it is also in the SNPs that one can find the variants which are linked to a rise in sensitivity to various diseases. Scientists have already identified twelve million SNPs, each of which exist in a minimum of two different forms in the worldwide population (www.hapmap.org/whatishapmap.html.en). A new research method has emerged in recent years, called Whole genome association, which consists of screening a genome in a systematic search for the different forms of these SNPs.

It is this method, used for the first time in Belgium, that the Liège researchers adopted, with the following hypothesis in their sights: there must exist differences in the form of these SNPs within the genome of people affected by Crohn’s disease that one does not find in healthy people. Cécile Libioulle was hired by the laboratory of animal genetics, led by Michel Georges, to supervise this research. A first scan carried out on 300,000 SNPs enabled the researchers to isolate three zones which seemed pretty different in terms of the genomes of ill and healthy people. The analysis of this screening was carried out using the technological support equipment at the National Genotype Centre (Génopole d'Evry) and the GIGA-Research laboratory at the University of Liège. This analysis revealed three areas of particular interest. The first is situated on Chromosome 1, the second on Chromosome 5 and the third on Chromosome 16. The latter in particular encouraged the Liège team as other researchers had already identified it as a zone effectively associated with Crohn’s disease. There was a setback for the morale of those involved concerning Chromosome 1: an American team made the same discovery and published more quickly. There was only Chromosome 5 left! Time to move quickly! The competitors would offer no gifts and information in the possession of the Liège researchers indicated that other publications were in the
pipeline, which would render worthless all the years of effort, in terms of international renown in any case.

**Chromosome 5 in detail**

Without wasting a moment the second phase of the work was started: it was necessary to focus on this precise zone of Chromosome 5, to sequence it and thus make the differences between the patients and the control group appear. It was in this way that our researchers were able to display several variants in the same genetic sequence, one of which is found in 20% of those afflicted by Crohn's disease and in only 12% of healthy people. This is statistically significant, given the large number of people tested. The discovery was submitted to the American 'PLOS Genetics' journal which, fully aware of the race between different scientific teams, decided to publish the article immediately on its internet site, before publication in its paper version. This took place last February.

Some people might be surprised by is the fact that this zone is what is termed a 'gene desert', a non encoding zone. However it is fitted out with components which regulate a number of neighbouring genes implicated in inflammatory mechanisms. Amongst these genes one notably finds an E2 prostoglandin receptor, a substance which is released in large quantities in cases of inflammation. Would it suffice to block this receptor to get rid of the inflammation and treat Crohn's disease? Unfortunately, it's not that simple. Studies on mice have provided contradictory evidence. Rodents in which the gene which encodes the prostoglandin receptor has been eliminated (knock out mice) developed severe colitis, an inflammation of the intestine, as in Crohn's disease. And, conversely, researchers have shown that stimulating the production of this receptor increases the inflammation. In short, everything is a question of a subtle equilibrium. In any case, the genetic mutation revealed by the ULg researchers, and which is present in 20% of Crohn's disease sufferers, seems to increase the expression of this prostoglandin receptor.

*In this diagram are represented the different genetic variants identified in the 5th chromosome by Liège researchers. The sequence coloured green is that which is most often found in patients suffering from Crohn's disease (in 20% of cases).*
The researchers can in any case now focus on this receptor to find a substance which would either modulate its expression, block it, or on the contrary stimulate it. The potential medicine which would result from this work would not be, or so it is widely supposed, a remedy adapted to all patients, but only to those for whom the illness is linked to an anomaly on this precise area of the genome. ‘Generally speaking’, explains Doctor Edouard Louis, ‘the molecular medicine is shaping up to be a made to measure medicine, because it adapts to the large genetic variety of the patients.’

Thus we are looking at a genetic profile which is very much more complex than had been first imagined. After having hitched up their databank to those of English and American researchers, thus multiplying by three the statistical force of their data, the ULg researchers now know that it is not just 4 or 5 genes that are connected to the development of Crohn's disease but 5 or even 10 times as many. Not every patient possesses all the sequences identified, but without doubt they have around ten of them and it is from this combination that the disease is born. That is not to say that it is necessary to work on the whole of the genes in this combination to obtain therapeutic effects. A single one would perhaps be enough, as it plays a central role in the molecular cascade that leads to the inflammatory illness.

The study carried out in Liège was one of a series of three studies. Pooling the results from these studies allowed more than 30 chromosome regions to be shown that were associated with Crohn's disease.

(read the account of this study published in June 2008 in Nature genetics).