Human Genome Project is 10: Where are we now?

March 2 2010

"It's hard to think back and remember how we worked then. We were scrabbling around in the dark," says Professor Mark McCarthy of the Oxford Centre for Diabetes, Endocrinology and Metabolism [OCDEM], recalling how research on the genetic causes of disease had to be carried out before the human genome was sequenced.

The first draft sequence of the human genome was announced at the White House 10 years ago this June by Bill Clinton, with the promise that it would lead to new ways to prevent, diagnose and treat disease.

Mark McCarthy, who is also at the Wellcome Trust Centre for Human Genetics [WTCHG], is the ideal person to explain what has happened since researchers got their hands on the DNA code and where we are now, 10 years on. Mark's research aims to identify genes involved in

diabetes and obesity, and he has been a leader in international collaborations to use the latest genotyping technology to advance our knowledge in this area.

Genetics of diabetes & obesity

One in ten people either has diabetes now or will get it in future, making it a global health challenge now and in the coming decades.

‘Not everyone exposed to our increasingly sedentary lifestyles becomes overweight,’ says Mark, ‘meaning obesity is the result of a combination of nature and nurture, the effects of genes and the environment.'

‘In diabetes we’ve found just short of 40 genes associated with an increased risk of the condition. For obesity, the number’s similar.’

He adds: ‘For diabetes, some of these genes are involved in cell cycle regulation and so may help in maintaining the cells in the pancreas that produce insulin. Perhaps some people are blessed with more or better islet cells in the pancreas.'

‘The evidence from the genes identified for obesity supports the involvement of the part of the brain involved with appetite and the feeling of satiety. It seems that subtle effects in the neural circuitry in the brain that controls what and how much we eat are involved here.’

For Mark, the importance of the human genome is clear: ‘We can trace all this progress back to the sequencing of the human genome. It was an overwhelmingly positive step which has stood the test of time.'

‘Genetics is one of the tools we can use to unlock the mysteries of disease. By identifying the bits of biological machinery involved, we can then use this to treat disease.’

Sequencing the genome

Before the Human Genome Project, researchers wanting to identify genes involved in disease processes would have to rely on laborious or haphazard techniques - ones that didn't reveal much detail or ones that gave no sense of the bigger picture.

Researchers might select one or two candidates genes that, at best guess, might be involved in disease ('many were looked at, many turned out not to be involved,' says Mark). Or scientists could look at broad genetic markers and see how often they occurred in families - perhaps they were linked or often inherited alongside causative genes for disease.

The human genome gave scientists the data they needed to systematically search for, identify, and determine the roles of crucial genes.

'The first human genome was a hugely influential and transformative step. It gave us a systematic way of looking at genes and heralded an era when genetics became 'big science' like astronomy and physics,' says Mark.

Once that first human genome had been sequenced, an obvious next step was to understand the genetic origin of variation between people - to find the 1 per cent of the genome that differed between people and where it was located. This could now be attempted in a comprehensive, systematic way across the 3 billion letters in the human DNA code.

Two international collaborative efforts involving many hundreds of scientists, the SNP Consortium and the International HapMap Project, mapped the locations of common single-letter changes in the DNA code of different people.

With that powerful knowledge in hand, it became possible to see if any

of these differences in individual genetic makeup could explain people's predisposition to common diseases like cancer, heart disease and diabetes.

'These individual differences were the logical place to start looking for differences in people's DNA that can lead to disease,' Mark explains.

'Three things came together at the same time, creating a 'perfect storm',' he says. 'The HapMap project told us where common variation was; improvements in the accuracy and cost of genotyping technology meant we could look at hundreds of thousands of locations on the genome; and there was the realisation that a few hundred samples were not enough. We needed larger-scale studies of thousands of people, or to combine multiple studies to give us reliable results.'

Scanning for genes and disease

The result was a new industry of big, international studies that scanned the whole genome of thousands of people using the latest technology. The studies checked the large number of sites where individual differences were now known to occur, to see whether common genetic variations could be associated with a particular disease or condition. And this huge effort has been very successful.

'Close to 1000 sites of genetic variation have now been associated with common diseases, such as diabetes, heart disease and cancer,' Mark says. 'That is way more than we might have expected and it has given us insights into many diseases.'

'At the same time, the results have also been surprising in that the variants we've identified don't explain more than a minority of the genetic basis for disease, and it has proved harder than we imagined to translate the results of these studies rapidly into new biological insights.'

The fact that, after all these studies, most of the genetic basis for common diseases still remains unknown has led to the concept of ‘missing heritability’. For example, despite almost 40 genes having been connected to diabetes and another 40 to obesity, in both cases the identified gene regions account for only around 10 per cent of the known risk that is inherited of developing these conditions.

‘There has been a lot of discussion about the 'missing heritability', comparing it to dark matter in physics - something that we know must be there but haven’t discovered yet,’ says Mark. ‘There are many possibilities for what makes up that missing heritability.’

The most likely, he says, is that less common or rare variants in the DNA code - which would not arise often enough to be picked up in genome-wide screens - have big impacts. Although rare, some of these variants would be connected with a much greater increased risk of disease over those found so far, and help to account for that missing heritability. And the search is now on.

New technology, new studies

Advances in sequencing technology mean it now should be easier to look for these rare sites and test them systematically for association with common diseases - cancer, heart disease, or even schizophrenia, for example.

Rather than scan the genome as before, checking individual locations (albeit many thousands of them), the aim is now to sequence all of the DNA for thousands of people - with and without disease - and find genetic differences linked with various conditions.

The 1000 Genomes Project is already sequencing the genomes of different people to provide a more complete understanding of sequence

variation between individuals than we have had so far. It's sequenced around 300 people so far. But as the technology improves almost month by month, even larger projects become possible.

'It took 10 years, $3 billion and many scientists around the world to sequence the first human genome,' Mark explains. 'Huge technological changes in the capacity to generate data means a handful of people can now generate a genome in a week for maybe $30-40,000. But these figures are almost out of date as soon as you write them down.'

'New studies that compare whole genome sequences should be able to pick up common and rare variants in a way that has not been possible to date. There is optimism that this will nail the missing heritability. The first efforts to do this on a large scale are now beginning or being planned.'

Mark McCarthy is part of a new transatlantic consortium of researchers that is setting out to sequence the whole genomes of 3000 people: 1500 with diabetes and 1500 without.'

'This will be a large effort. 3000 people is ten times the size of any dataset currently available,' he says. 'We calculate that the project will allow us to detect variants that are present in 1 in 100 people that are associated with diabetes.'

'Previously, we've done very well at detecting variants present at a level of 1 in 10 people, but there will be many more variants present at lower frequencies that that.'How soon this new knowledge will translate into new genetics-based medicines?'

'That's the big open-ended question. It is very easy to talk about possibilities for new treatments and underestimate the time it takes. But we can take hope where rare mutations causing diabetes have been

identified - mutations that may be responsible for perhaps 2 per cent of diabetes cases. It has been possible to use that information to come up with new approaches for diagnosis and treatment for these rare types of diabetes.'

'It is not out of bounds that we could do something similar with the information we obtain in these new sequencing studies,' says Mark.

More information: www.ornl.gov/sci/techresources … an_Genome/home.shtml

Provided by Oxford University