Human Epigenome Project - Up and Running

Jane Bradbury

The Wellcome Trust Sanger Institute and Epigenomics AG recently announced the launch of the Human Epigenome Project - a 5-year project to map the sites of DNA methylation throughout the human genome.

Citation: Bradbury J (2003) Human Epigenome Project—Up and Running. PLoS Biol 1(3): e82. doi:10.1371/journal.pbio.0000082

Published: December 22, 2003

Abbreviations: HEP, Human Epigenome Project; MHC, major histocompatibility complex

Jane Bradbury is a freelance science news writer based in Cambridge, United Kingdom.

Epigenomics is one of the many ‘omics’ that is being talked about in the wake of the Human Genome Project. But what is an epigenome, and why have the Wellcome Trust Sanger Institute (Hinxton, United Kingdom) and Epigenomics AG (Berlin, Germany) recently announced the launch of the Human Epigenome Project (HEP), a five-year undertaking during which DNA methylation sites throughout the human genome will be mapped? The HEP is the brainchild of immunogeneticist Stephan Beck of the Sanger Institute and Alexander Olek, chief executive officer of Epigenomics AG. The Human Genome Project, explains Olek, ‘provided the blueprint for life, but the epigenome will tell us how this whole thing gets executed’, what determines when and where genes are switched on and off to produce a person. And knowing more about the human epigenome may provide clues to what goes wrong in cancer and other diseases.

What Is Epigenetics?

Most people have a fair idea of what is meant by genetics. They know that characteristics such as eye colour are specified by the DNA sequence within their genome. But not everything is that simple. For example, genetically identical twins can be very different. ‘One might be normal, while the other is autistic’, explains chromatin researcher David Allis (Rockefeller University, New York, United States). ‘We can't explain that on the basis of pure genetics because the DNA is identical. Something else must be at play’.

The Human Genome Project … ‘provided the blueprint for life, but the epigenome will tell us how this whole thing gets executed’.

That ‘something else’ is chemical modifications of genes that are heritable from one cell generation to the next and that affect gene expression but do not alter the DNA sequence. Epigenetic modifications can affect the DNA itself or the proteins that package the DNA into chromatin. Developmental geneticist Wolf Reik (The Babraham Institute, Cambridge, United Kingdom) describes these modifications as ‘red and green traffic lights that are superimposed on top of the genome to tell the genes whether they should be active or inactive’. For individual cells, this code on top of a code helps to determine whether a cell is a blood cell, a fat cell, or

something else. And in the case of monozygotic twins, unexpected differences may result from chance variations in this superimposed code.

The study of these modifications—what they are, how they are laid down, and the processes that they control—is the field of research known as epigenetics. An epigenome is the description of these modifications across the whole genome, but unlike the genome DNA sequence, each organism has multiple epigenomes—for example, in different cell types—that may change during its lifetime in response to environmental cues.

DNA Methylation

Methylated DNA was the first epigenetic mark to be discovered. That this epigenetic modification is important is suggested by the waves of demethylation and de novo (new) methylation that occur at specific stages during the development of an animal from a fertilised egg. Soon after fertilisation, there is a massive active loss of methylation from the paternal genome, explains Reik. At the same time, the maternal genome loses some methyl groups through passive demethylation. ‘We think this process erases all the epigenetic memory in the gametes [sperm and egg cells] except for some special imprinted genes whose expression depends on whether they are of maternal or parental origin’, says Reik. Later in development, widespread de novo DNA methylation occurs in a process known as reprogramming. De novo methylation rarely occurs after an early developmental event known as gastrulation, except in cancer cells where special unmethylated regions of the genome, known as CpG islands, often become methylated.

‘People generally agree that DNA methylation is important for imprinting and reprogramming’, says Reik, ‘but not everyone agrees that it plays a role in DNA activation and inactivation during development’, a role first proposed more than 25 years ago. ‘People are still hesitant about saying that DNA methylation is important in making sure the right genes are expressed in the right cells at the right time because until recently we have had relatively few good examples’, says DNA methylation expert Adrian Bird (The Wellcome Trust Centre for Cell Biology, Edinburgh, United Kingdom). Many people have reported correlations between methylation states and gene activity but, says Reik, ‘you can look at such correlations until the cows come home and it tells you nothing about causality’.

The HEP Pilot

Even given the doubts about the exact role of DNA methylation, Bird, Allis, and Reik believe that mapping methylation patterns across the human genome, which is what the HEP plans to do, is a worthwhile albeit enormous undertaking. ‘Knowing the methylation sequence’, says Bird, ‘will be an essential backdrop to future research on DNA methylation and its biological effects’. Allis agrees that DNA methylation is an important part of epigenetic marking, but ‘although mapping DNA methylation patterns is the logical and good place to start, by its lonesome it won't explain the whole epigenetic phenomenon’.

Beck and Olek began thinking about mapping the human epigenome more than five years ago. Even as long ago as 1998, explains Olek, ‘it was pretty natural to be thinking about what would come after the Human Genome Project, and Stephan believed that it had to be methylation sequencing because that has the potential to tell us about the hundreds of genomes we really have’.

In October 2000, the Human Epigenome Consortium (the Sanger Institute, Epigenomics AG, and the Centre National de Génotypage in Evry, France) started a European Union-funded pilot project to map the methylation sites within the major histocompatibility complex (MHC) region in seven different human tissues. Rather than look at all the DNA methylation sites throughout the entire MHC region, the consortium set out to look at the 150 expressed genes within this region, explains Beck. For each gene, the scientists chose two areas to analyse for methylation, each about 500 basepairs long. ‘One window was in what we thought was the promoter region’, says Beck, ‘and the other corresponded to the most CpG-rich region in the gene’. The promoter regions are the places where the elements that control gene expression are often located, and DNA methylation occurs at cytosine residues within CpG motifs, hence the choice of CpG-rich regions.

The methylation status of more than 100, 000 sites was determined during the three-year pilot study and the results analysed to show where there were methylation differences in the MHC between different tissues. ‘We found major methylation differences between loci and between tissues’, says Beck, ‘and we are particularly interested in what we call methylation variable positions (MVPs), which we believe will advance our ability to understand and diagnose human disease’. Olek echoes Beck's satisfaction with the results to date. ‘We can see that there is hidden treasure in the data that needs mining, and that treasure is obviously going to get much bigger when we sequence the whole thing’, he says.

On to the Full HEP: A Public/Private Initiative

That endeavour—the mapping of DNA methylation sites in all 30,000 human genes in around 200 samples—is now underway. ‘These will be very carefully chosen’, says Olek, ‘with the help of additional academic centres who we hope will join our consortium’. Sample preparation and amplification will be done at Epigenomics AG, the Sanger Institute will do the sequencing and raw data analysis, and then the data will be jointly evaluated and mined, ‘a huge collaborative experiment’, says Olek.

The Wellcome Trust and Epigenomics AG are jointly funding phase I of the HEP, thus avoiding a rerun of the situation that occurred in the Human Genome Project, in which a commercial company and a public effort ended up competing with each other. An agreement has been drawn up that ensures that both partners benefit from the collaboration, and all the data generated will be publicly available through the Internet in accordance with the consortium's data release policy.

The mapping of DNA methylation sites in all 30,000 human genes in around 200 samples is now underway.

Histone Modifications and Epigenetic Landscaping

DNA methylation is not the only epigenetic mark. There are also a staggering number of acetylation and methylation marks on the histones—the proteins that bind DNA to form chromatin. Histone modifications alter the chromatin structure and thus regulate gene expression, says Allis, ‘and somehow the cell exploits each and every one of these in a meaningful way’. Recently, a link between DNA methylation and histone methylation was uncovered, a finding that Reik says has ‘galvanised the field’. In Neurospora and Arabidopsis, he explains, there is now good evidence that histone modification puts down some sort of mark that is read by a series of binding proteins, including the enzymes that methylate DNA. How the positional specificity of the first mark is determined remained a mystery until the end of 2002, when the first reports appeared suggesting that in Arabidopsis small RNAs might be involved. The small RNAs are thought to bind to ‘complementary’ sequences in DNA and somehow target the region for epigenetic modification. Whether similar mechanisms exist in animals remains to be seen.

To answer this and other epigenetic puzzles, scientists are organising themselves into collaborative networks to maximise progress. In Europe, for example, Thomas Jenuwein (Research Institute of Molecular Pathology, Vienna, Austria) is in negotiation with the European Commission over the establishment of a Network of Excellence in epigenetic research. This 'Epigenome Network', which will focus on epigenetic mechanisms, 'will allow us to build a long-lasting platform to bring together colleagues working on epigenetics in Europe and make this area of research very strong', says Jenuwein. Negotiations with the European Commission for an anticipated €12 million of funding over five years should be completed soon, and, if all goes well, the Network should start work in Spring 2004. The Network will not itself fund any epigenome mapping projects like HEP, says Jenuwein, but will build strong links with scientists involved in mapping the 'epigenetic landscape', the pattern of DNA methylations and histone modifications.

Just an Academic Exercise?

Epigenetics is a fascinating phenomenon, but why are funding bodies keen to fund expensive collaborative enterprises like the HEP and the Epigenome Network? One hope is that an understanding of how cells execute their genomes normally will provide important clues about what goes wrong in diseases such as cancer, in which, says Bird, 'methylation has gone a bit wonky in various respects. In particular, CpG islands, regions of the genome that are normally nonmethylated, get methylated, and in some cases this seems to shut off tumour suppressor expression'. Loss of tumour suppressor function can remove the controls that normally restrict cell division, leading to the unrestrained cell growth that is characteristic of cancer. There are also alterations in histone acetylation patterns in cancer, and cataloguing these epigenetic changes in different tumours may provide a way to tag cancers to help clinicians to choose the best therapeutic regimen for individual patients. Epigenetic changes may also provide scope for new therapeutic approaches, says Allis. If a gene is inappropriately switched off during cancer development, he explains, maybe if we could de-silence it, we could reverse tumour development'. Bird agrees that reversing epigenetic marks in this way is an experiment that has to be done. But, he asks, is epigenetics the Achilles heel of cancer, or are there too many genetic changes that are hard-wired to be able to reverse it simply by interfering with epigenetics?'

Recently, a link between DNA methylation and histone methylation was uncovered, a finding that [Wolf] Reik says has 'galvanised the field'.

A better understanding of epigenetics might also improve the efficiency of cloning animals from somatic cells (reproductive cloning). It is very hard to clone animals from somatic cells and, says Bird, ‘the more you look at the surviving cloned animals in molecular terms, the odder they look’. Reik suggests that this could indicate that epigenetic reprogramming—the process needed to turn a somatic cell back into a totipotent cell—is very inefficient. ‘We have some evidence that there is a link between the success of epigenetic reprogramming and that of cloning, and we really have to understand the reprogramming process to get reproductive cloning to work', says Reik. ‘At the moment the whole process is a black box’. Because of this, many scientists argue that human reproductive cloning, leaving aside the ethical concerns about it, simply isn't a safe option at present, a concern reflected in the recent call by 60 scientific academies for a ban on its development.

Epigenetics then, both at the level of unravelling mechanisms and mapping epigenomes and at the applied level, is an exciting area of research. A far cry from the situation 10 years ago, says Reik, ‘when the field had the smell of something esoteric that no one really understood and whose significance was not clear. Now it is one of the mainstream exciting areas of post-genomic biology’, he concludes, and one which will hopefully yield to the might of large consortia like the Epigenome Network and the HEP.

Where to Find Out More

Information on the project, publications, and up-to-date data releases are posted at http://www.epigenome.org
.

The DNA Methylation Society homepage provides more details on DNA methylation at http://dnamethsoc.server101.com
.