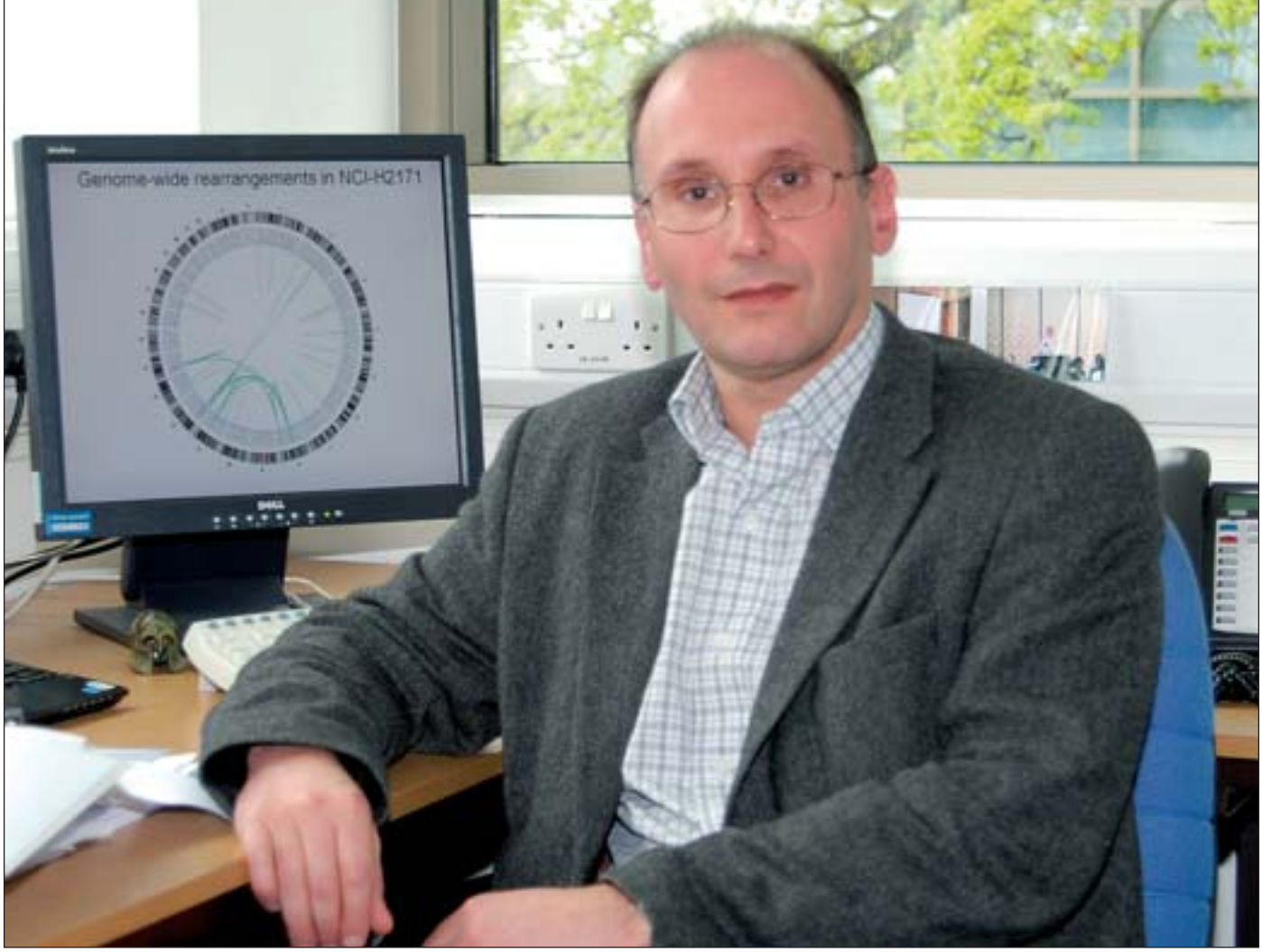

Photos (2): Wellcome Trust Sanger Institute

A conversation with Michael Stratton, Hinxton

"Drivers and Passengers"

Cancer geneticist Michael Stratton talks about how high throughput genomics helps identify mutational patterns of cancer susceptibility genes and how, in addition, they reveal the "histories" of individual cancers.

LT: When did you start thinking about how to use the human genome sequence for cancer research?

Stratton: About 1998 or 1999. Our premise, that we know is true, is that cancers are due to abnormalities in DNA. Over the last 30 years, the identification of those key genes that make a cell turn into a cancer cell has been an essential aim of cancer research. It has been successful. Cancer genes were found using several strategies, such as looking down the microscope into the nuclei of cells, seeing chromosome rearrangements that turned out to be the reason for particular cancers such as leukaemias and lymphomas. But these approaches were running out, they were no longer delivering cancer genes. I thought that on the basis of the human sequence we could apply new strategies, essentially, in the long term re-sequencing the genome of a cancer, find all the differences from the normal human genome sequence from that same person and thus extract the full set of somatic mutations in that cancer genome

How many genes do you expect are involved in transforming a cell into a malignant cancer cell?

Stratton: There is a discourse in the literature that perhaps in a common cancer, like breast or colon cancer, five or seven genes in each cancer in each individual are abnormal, driving that cancer but this is very indirect evidence. Actually, we don't know how many genes drive an individual cancer.

"We sequenced kinase genes in each of 200 cancers: lung cancers, breast cancers, colorectal cancers, gliomas and so on. That was the first time that we were able to look at cancer genomes in this way."

How did you approach your vision?

Stratton: Until last year or so, the sequencing technology available didn't allow us to implement the vision in the way I just described. So what we did, and what similarly the Americans also do, is we started with key experiments. We sequenced the coding parts of 500 protein kinases that sum up to 1.3 million base pairs. Kinases are enzymes involved in cell signalling. We sequenced those genes in each of 200 cancers: lung cancers, breast cancers, colorectal cancers, gliomas and so on. That was the first time that we were able to look at cancer genomes in this way. We found one to two somatic mutations per million bases. So an average cancer genome contains about 5,000 of these individual somatic mutations distributed throughout the genome.

Did you identify new kinases involved in cancer?

Stratton: We identified new key cancer genes, for example BRAF gene. This gene

encodes a kinase which we knew existed, but we didn't know that the gene was mutated in cancer. In 2002, when we found that gene to be involved in cancer, there were no clues about this gene, neither biological nor physical clues. None of the previous strategies would have drawn us to this gene but by the systematic gene sequencing approach we found BRAF to be mutated in about 70% of malignant melanoma. In addition, we found evidence that there were many more kinase mutations contributing to human cancer, but these were less common.

Did your discovery lead to advances in drug development?

Stratton: Yes, they spawned a huge effort in drug development in the academic and pharmaceutical sectors, particularly targetting BRAF. The first drugs are now beginning to enter Phase I trials.

Meaning that the industry is amazingly quick off the mark.

Stratton: Mutations like in the BRAF gene are what the drug development people like. It's easy to develop a drug against an enzyme and it's easier to inactivate an enzyme than to activate it. Mutations in the gene for the enzyme BRAF switch the gene on permanently. The other reason why our finding was taken up by the industry is a precedent, the development of Gleevec.

That's the name for an inhibitor of tyrosine kinases.

Stratton: This drug inhibits the kinase Abl which was known to be activated in chronic myelogenous leukaemia, called CML, and other malignancies. In Phase I drug tests of Gleevec in CML most patients went into remission. The evidence was so strong that the FDA approved Gleevec on the basis of these Phase I trials. It's now a standard treatment for CML. It was the first time that molecular biology and genetics of cancer had led to a new drug and the new drug had an absolutely dramatic effect. I can't communicate to you how much this discovery changed the world in cancer research.

In which way?

Stratton: Previously, we had a lot of good geneticists and molecular biologists in cancer research and we had plenty of good oncologists. These two groups rarely talked. Suddenly, both of them have realised that their worlds come together, that science has generated a drug, a treatment.

"It was the first time that molecular biology and genetics of cancer had led to a new drug and the new drug had an absolutely dramatic effect. I can't communicate to you how much this discovery changed the world in cancer research."

How many cancer driving genes are now known?

Stratton: Roughly 400 genes have been found, which cause cells to become a cancer cell when they are mutated in one way or another. That number is worth thinking a bit about because there are roughly 20,000 genes in the genome. That means that 2% of the protein-encoding genes in

the genome are mutated in one or another of the some 100 types of cancers.

With the whole-genome approach you'll not only identify cancer genes but all mutations a cell has acquired throughout its lifetime. How will you distinguish between them?

Stratton: Essentially, when it divides, every cell in the body occasionally gets mutations. Most don't do any harm; we call them passenger mutations. To sort out the difference between the drivers and the passengers we have to look at the genome of many cancers. In general, the passenger mutations will be distributed randomly throughout the genome, but the drivers will occur in specific genes in a particular cancer type. You would expect the same driver genes to be mutated in multiple different samples. We identified drivers and passengers. Our data suggests that we had about 100 cancer causing mutations in the 200 cancer samples we looked at through the kinases but many of them were found only once or twice in a particular gene. We think we have 60 to 100 new cancer genes. Many new cancer genes are mutated infrequently in any particular type of cancer. That's the main message. On the one hand, that is quite daunting because if there are a lot of cancer genes how can we make drugs against all of them? But on the other hand, we're now beginning to see the lie of the land. There is a lot more to be found.

Could the analysis of the passengers lead to any new conclusions?

Stratton: That's a good question. Why should anybody be interested in these mutations? They're just there, they don't make any trouble. However, the passenger mutations reflect the exposures through which that cancer has gone.

You mean you could deduce from the sorts of passenger mutations in cancer whether the patient smoked or drank too much alcohol?

Stratton: Yes. Cells are mutating all the time as they grow. But sometimes mutations are caused by exogenous factors such as smoking or radiation. You can see the imprint of those exposures in the DNA because radiation causes a different pattern of mutations than tobacco smoke. It's like an archaeological record of the exposures the cancer has gone through.

"Radiation causes a different pattern of mutations than tobacco smoke. It's like an archaeological record of the exposures the cancer has gone through."

What surprised us was that we found a particular type of mutation pattern in colorectal cancer, although we are not sure what it is. In breast cancer we found a completely different pattern. We looked in the literature but we can't find any explanation of it at all. We asked colleagues who are interested in the processes of mutations – but they don't know, either. The signature left in breast cancer cells must be of something that has happened to them in the past.

When considering the history of a cancer, wouldn't it also be worth sequencing the genomes of benign tumours that later develop into malignant tumours?

Stratton: If you make the assumption that the benign tumour is the early stage in the development of a malignant one, you may get an idea about the succession of the changes in the genes. Some benign tumours, however, never become malignant tumours. For example, there is a kind of ovarian tumour called borderline ovarian cancer. This tumour has to be removed, though it is generally a dead end of cancerogenesis, it hardly ever transforms into malignant cancer. Why not? It's fascinating to learn about why these proliferating cells can't go any further.

You're currently working with cell lines. Do you expect to see differences between cultured cells and fresh tumour tissue?

Stratton: We have done a lot of work with cancer cell lines. They are easy to use, one can get large amounts of material for analysis from them and one can interrogate them biologically. On the other hand, they have been growing in culture for long periods, sometimes decades, and during that time they may have acquired additional mutations. Thus we would expect to see some differences between cultured cells and primary tissues. However, the current evidence is that there may not be as many as people have speculated in the past.

Now that you have finished your initial experiments, what's coming next?

Stratton: We are still working our way through our big vision. Others are doing similar sorts of studies; for example the American research groups working on The Cancer Genome Atlas – TCGA. Thus far, the sorts of changes in the DNA that we've been seeing have been of two types. In the first type of study we looked at point mutations. Secondly, we looked at the copy number of genes. However, it has been very difficult to look at translocations, rearrangements of DNA. Those have been quite hard to find but they are clearly very important and they occur frequently in lymphomas and leukaemias. Until very recently, we had not known about these rearranged genes in common cancers of the breast, colon and lung. That produced a big discussion amongst cancer geneticists. One side said there is something special about genes that are rearranged in this way, which means that they can only contribute to cancer in particular types of cells, such as blood cells and lymphoid cells but they can't work in the epithelial cells,

Michael Stratton

is Head of the Cancer Genome Project at the Wellcome Trust Sanger Institute in Hinxton and Professor of Cancer Genetics at the Institute of Cancer Research, London. He obtained a Ph.D. in the Molecular Biology of Cancer at the Institute of Cancer Research in 1989. His past work has mainly been directed towards the study of cancer susceptibility genes. He leads the group that mapped and identified the breast cancer susceptibility gene BRCA2, and he has also identified susceptibility genes for skin, testicular, colorectal, thyroid, and childhood cancers.

The Cancer Genome Project started at the Wellcome Trust Sanger Institute in 2000. It aims to conduct high-throughput, systematic genome-wide searches for somatic mutations in human cancer in order to identify new cancer genes.

which cause common cancers. The other group said, "we just haven't found them."

Who was right?

Stratton: Over two years, one or two of these genes have been found in common cancers. 70% of prostate cancers carry translocations or rearrangements in a special gene. So now, from apparently not happening in epithelial cancers, we now know that the most frequently occurring rearranged cancer gene is in an epithelial cancers, prostate cancer. There are likely to be many more. Rearrangements in a particular family of genes have also been found in lung cancer, particularly adenocarcinoma.

Which one? More details please! How have they been identified? How will you identify rearrangements?

Stratton: Until now rearrangements in common epithelial cancers have been found in several ways. However, with the newly arrived, second generation sequencing technologies we can directly look for rearrangements in the DNA of cancer genomes. One approach is to sequence both ends of a large number of pieces of DNA (paired-end sequencing). The sequences from both ends of each piece are mapped back to the genome. If they map back close to each other, as they should, then one assumes that the piece of DNA between them is normal. However, if the two ends map back, for example, to different chromosomes there must be a rearrangement between them which we can then further characterise.

"Today, you get 100 times more sequence in the same time for the same cost, than two years ago. By the end of the year it will probably go up to 1,000 times."

What does it cost to sequence the genomes of one cancer?

Stratton: To completely sequence the genome of a single cancer sample at sufficient coverage currently costs about $100,000. Of course one then has to sequence the normal DNA from the same person to find all the somatic changes that are particular to the cancer, so double that.

That's a substantial amount. Is that the reason that in Toronto recently several research groups from all over the world formed the International Cancer Genome Consortium (ICGC)?

Stratton: If we are to do the job properly, we need to sequence hundreds of cancers of a particular type and there are 100-200 types of cancer. The current technologies are simply too expensive. However, the ICGC predicted that in a few years it may be possible to do several hundred examples of a single cancer type for approximately 20 million dollars. Anybody who is part of this consortium is a funder. They need to think about whether they are willing to provide 20 million dollars over a five year period. Willingness to make that commitment means you're into the consortium. This will be the first phase. Over the next two to three years we will have six to ten projects going, the ones that are working out the problems, sorting the issues out. It will be easier to get more money from other, smaller funding organisations if the costs of sequencing come down.

Today, you get 100 times more sequence in the same time for the same cost, than two years ago. By the end of the year it will probably go up to 1,000 times. From our machines we are getting 4 gigabases per run in about a week. Prices are going to change as the sequencing improves. The more we show the industry that there's much more sequencing to be done, the more they are going to work on their technologies and that will reduce the prices.

Is The Cancer Genome Atlas (TCGA), the American initiative, involved in that consortium?

Stratton: Yes, they are. The ICGC is set up in a similar way to the human genome project. It's the largest sequencing project at the moment to have been announced. It's a huge amount; more than the 1,000 genomes project. We want to sequence 500 examples each of 50 cancer types (25,000 cancers in total) and we have to sequence the normal DNA from each person as well. That's 50,000 genomes, which will be sequenced to a 30-fold coverage.

Let's look into our crystal ball. When will we have the first cancer genomes?

Stratton: We'll be starting slowly. Some of us, one or two groups in the world, will work on one cancer genome, at 30-fold coverage. I suspect that by the end of 2008 there will be a couple of cancer genomes that will have been sequenced and described. By the end of 2009 that will have gone up to a few, perhaps 10. What we can say with confidence is that there are a lot more genes to be found.

INTERVIEW: KARIN HOLLRICHER