

# The genomics of cancer

Many cancers are associated with changes to our genetic material, DNA. These may be small, single unit substitutions, large rearrangements such as deletions or duplications of a part of the DNA sequence, or various other forms of mutations. Although the smaller substitutions have been more intensively studied, **Dr Lixing Yang**, of the University of Chicago, focuses on uncovering changes at the larger end of the spectrum. His work suggests certain key variations may be connected to particular forms of the disease.

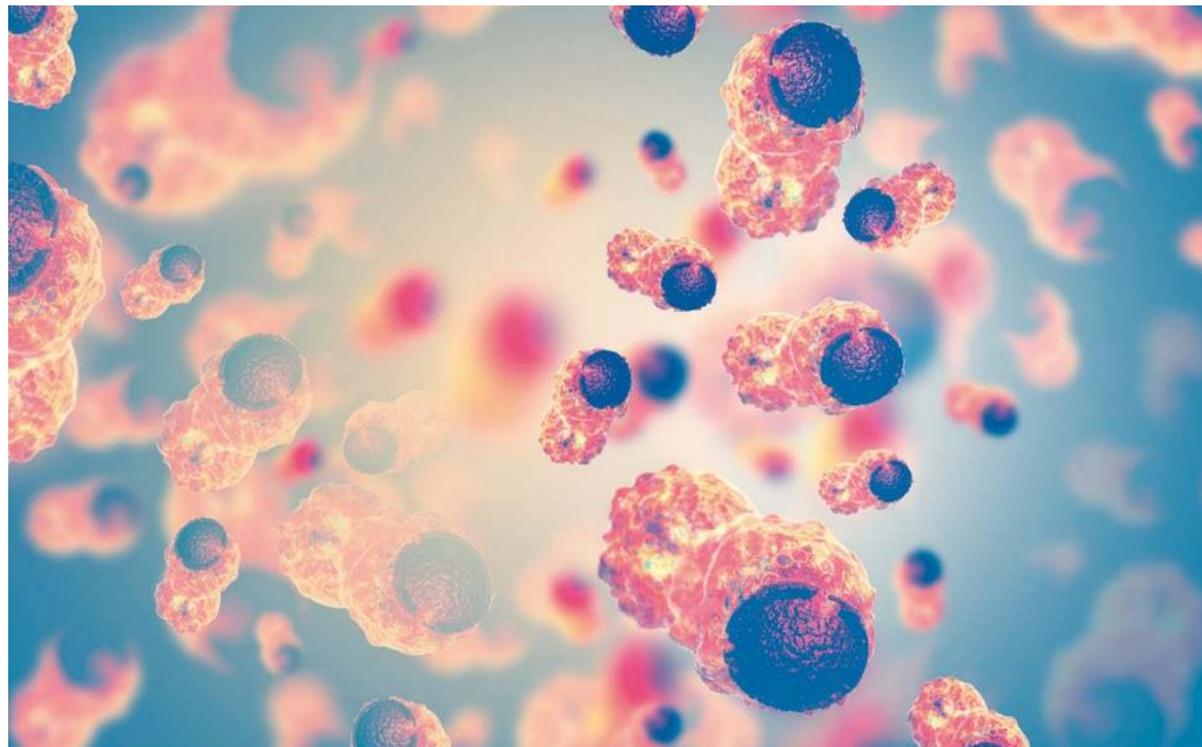
**W**hen we hear the term 'genetic disease', we probably think of something like cystic fibrosis, or Huntington's Disease. But, increasingly, cancers are also being linked to genetic causes, whether an inherited trait found in every cell of the body, or a change occurring in just the cells of a tumour itself.

## REARRANGING THE GENOME

Genes are held on chromosomes, aggregates of DNA and protein found in almost every cell of the body. DNA is the blueprint of most living organisms. Our DNA is made up of a sequence of paired sub-units known as nucleotides or bases, denoted by the letters A, C, G and T. DNA changes can be as

small as a single substitution: a T for an A, say, or a C for a G; or they can be as large as an entire portion of one chromosome – thousands or millions of As, Cs, Gs and Ts – being transferred to another. The larger changes may affect an order of magnitude more DNA than single nucleotide variations, yet the functional impact of such 'genomic rearrangements' has until now been largely unexplored.

Genomic (or 'structural') rearrangements are commonly generated during DNA replication, or when the cellular machinery that repairs damaged DNA is not functioning correctly. They may take many forms and arise from multiple different events, such as deletions,



duplications, insertions, inversions or translocations of DNA within and between chromosomes.

Recent advances in technology now make it cheaper, quicker and easier than ever before to determine the entire sequence of the DNA held in the cells of an organism. This 'whole genome sequencing' has revolutionised many areas of biology – and cancer biology is no exception. Dr Yang's research focuses on using whole genome data to discover larger scale structural rearrangements of our genetic material that may cause, or contribute to, the development of cancer, and to determine the relative contributions of these different kinds of rearrangements in different kinds of cancers.

## MEERKAT TO THE RESCUE

The human genome is over three billion base pairs long – so searching for genomic rearrangements in its sequence is no easy matter. For this reason, Dr Yang has developed a novel computer program to facilitate his quest. The algorithm, Meerkat, can both identify more complex rearrangements than available methods and ensure high levels of accuracy in pinpointing the location and nature of structural rearrangements of the genome. Dr Yang has already used Meerkat to catalogue structural variations found in over a thousand tumours from patients with dozens of different types of cancer.

His research has shown that in tumours, two forms of DNA damage – breaks in both strands of the DNA double helix and errors in DNA replication – drive the majority of events leading to genomic rearrangements. For deletion-type structural changes, for instance, around 20% result from complex errors occurring when DNA is replicated.

Focusing on particular types of cancer, Dr Yang has found that in patients with a rare form of kidney cancer, chromophobe renal cell carcinoma, one gene, known as *TERT*, is sometimes activated. *TERT* is involved in cell proliferation and is essential to keep the chromosomes stable after many cycles of cell division. Dr Yang discovered



A DNA double helix

that 12% of these patients display genomic rearrangements near a DNA region known as a promoter, which controls the activity of the *TERT* gene. In glioblastoma multiforme, an aggressive form of brain cancer, a more complex picture is emerging: structural changes of many different kinds can cause the activation of several key oncogenes

## Increasingly, cancers are being linked to genetic causes

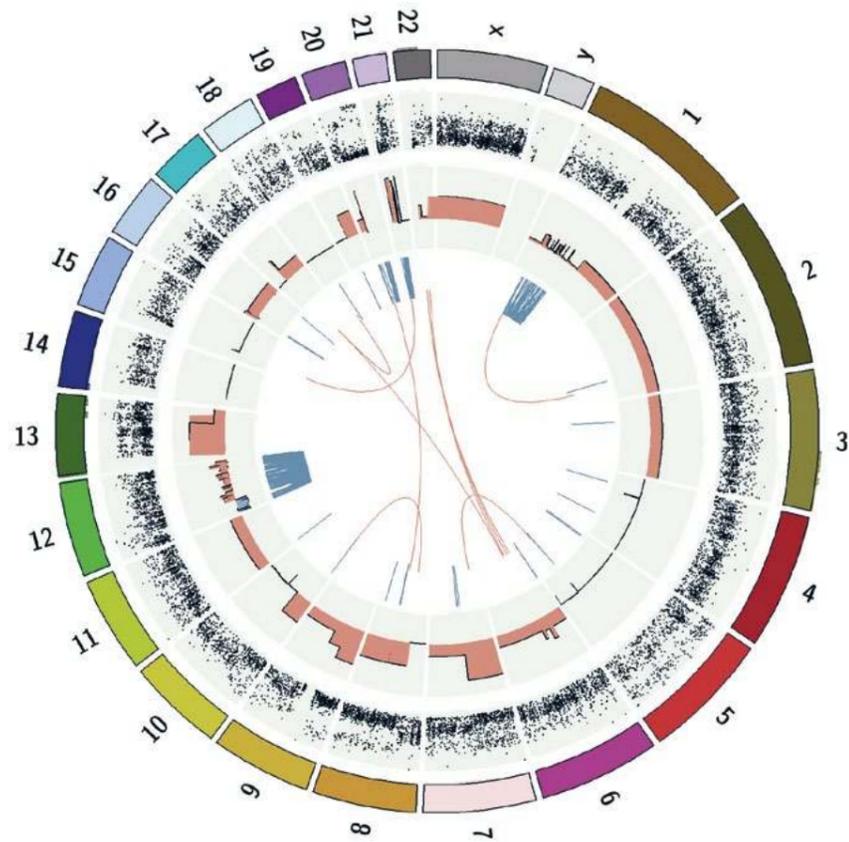
and inactivation of important tumour suppressors. In fact, says Dr Yang, his research has shown that the main

functional consequence of highly clustered genomic rearrangements in cancer is the switching on of oncogenes.

## A COMPLEX PICTURE

Dr Yang's research highlights just how intricate the interacting factors influencing DNA changes can be, with multiple mechanisms of mutation sometimes acting to cause rearrangements within one gene, while at the same time multiple driving forces acting on completely separate parts of the genome can together lead to an increased risk of cancer. It seems common that a positive feedback loop is present, where existing mutations

The genome view of somatic alterations in a colorectal cancer patient. From the outer rings to the inner ones: chromosomes 1 to Y, base substitutions, copy number changes (red - amplifications, blue - losses), intra-chromosomal rearrangements, inter-chromosomal rearrangements



in genes involved in DNA replication or repair result in an increased likelihood of further mutations including large structural rearrangements. Dr Yang has shown that almost all cancer patients have at least one alteration in a gene involved in DNA replication or repair, and almost half of these are themselves the result of structural rearrangements to the genome.

### Almost all cancer patients have at least one alteration in a gene involved in DNA replication or repair

While whole-genome-sequencing data, which offers unprecedented accuracy and resolution, is the 'holy grail' for researchers trying to identify DNA rearrangements, the time and expense needed to generate it remains a constraint. Much more widely available are so-called 'whole exome sequences' focusing only on the DNA regions thought to code for protein – known as 'exons.' The whole

exome sequencing data are typically only used to discover single base substitutions. Dr Yang has also improved his approach to scour exome data for structural rearrangements to re-use the vast amount of existing data, which has yielded further valuable findings.

On the other hand, although the greatest interest in structural rearrangements so far has been in those within genes themselves, there is growing evidence that changes in DNA regions previously considered 'non-coding' may also play an important, though perhaps more subtle, role. These changes might only be picked up through the whole genome

approach in which Dr Yang specialises. Dr Yang is quick to point out that this is an emerging field, and an inexact science. "The rearrangements we observe" he comments, "are a snapshot ... of the alterations in cancer genomes."

His research group at the University of Chicago is, he says, "battling on the frontiers of precision medicine." Precise characterisation of the mechanisms and nature of genomic rearrangements is crucial: by understanding the genetic mechanisms underlying tumour development, we come one step closer to discovering new drug targets, and ultimately, to helping patients.



# Behind the Bench

## Dr Lixing Yang

E: [lixingyang@uchicago.edu](mailto:lixingyang@uchicago.edu) T: +1 773 834 2948 W: <http://yanglab.me>

### Research Objectives

The goal of Dr Yang's research is to understand the roles of genomic rearrangements in human cancer by development and application of novel computational methods. In particular, he focuses on analysis of whole-genome sequencing data, which offers unprecedented accuracy and resolution in identifying such rearrangements.

### Funding

National Cancer Institute (NCI)

### Collaborators

- The Cancer Genome Atlas (TCGA) Research Network
- International Cancer Genome Consortium (ICGC)

### Bio

Dr Yang completed his PhD in plant genome evolution at the University of Georgia. He joined the Park lab at Harvard Medical School in 2010 before starting his own research group at the University of Chicago in January 2017, continuing his work on cancer genomics.

### Contact

Lixing Yang, PhD  
Assistant Professor  
Ben May Department for Cancer Research  
Department of Human Genetics  
Institute for Genomics & Systems Biology  
Comprehensive Cancer Center  
The University of Chicago  
929 E. 57th Street, GCIS W432  
Chicago, IL 60637  
USA

## Q&A

### How did you first become interested in this area of cancer research?

Cancer is a disease caused mostly by genetic and epigenetic alterations. Cancer research not only can help millions of patients, but can also advance our understanding on many fundamental processes of biology including organ development, various signalling pathways, metabolism, cell-cell interactions, ageing, immunology, and many others.

### How do structural rearrangements of the genome cause cancer?

The best-known cases are gene fusions – two different genes normally far away are fused into one new gene and the new gene performs a novel function in the cells. There are certainly other ways such as the relocation of regulatory

elements altering the gene functions that they target.

### What are the easiest types of rearrangements to identify? And the hardest?

The type doesn't matter too much; what matters is where they are. There are a lot of DNA fragments that have been repeated many times in the human genome and some copies are completely identical. The ones in such regions are the hardest to identify. The situation is more like being given two pictures: The Empire State Building and a random single family house. Which one is easier to identify?

### Does your research suggest that many cancers are heritable? Or that they result from mutations occurring during our lifetime? Or both?

There are certainly some inherited variants which increase the risk of cancer,

but most cancers are not heritable. There are damages occurring to DNA all the time and the vast majority are repaired. Only a very small number accumulate during each cell cycle. Over the years, more and more mutations accumulate in different cells. Once the mutations reach a certain combination, cancer starts to develop.

### How do you think your findings could be translated into treatments for cancer?

Many targeted therapies for cancer block protein functions of fusion genes. If we discover more rearrangement events that are responsible for tumourigenesis, some of the genes involved in the rearrangements could be good drug targets.