

The DNA detectives hunting the causes of cancer

September 25 2018, by Kat Arney



Halfway up a hill overlooking the Great Rift Valley in western Kenya are two graves. One of them is a few years old now, bristling with bushy shrubs stretching bright green leaves towards a cloudless sky. The other is a freshly dug bed of rough red dirt planted with a white wooden cross. They are the final resting places of Emily's mother and father, who died

within four years of each other.

Still a young woman, Emily now looks after her family's rural homestead near Iten – a town famed for churning out long-distance runners and playing host to Mo Farah's training camps. We reach it by driving through urban sprawl and out into the hills, passing a seemingly endless stream of impossibly fit athletes pounding the roadside paths.

Emily is busy cooking lunch when we arrive. Her kitchen is a small straw-capped mud hut built in the traditional style, similar to the other buildings that make up the homestead, with smoke pouring out of the door from an open fire and chickens scratching in the dirt nearby. It seems idyllic, but there's a killer on the loose around here, and we've come to track it down.

That killer is squamous cell oesophageal carcinoma – one of the two main forms of oesophageal [cancer](#), which starts from the cells lining the oesophagus. Cases started piling up more than 60 years ago in South Africa, when a doctor working in the Transkei territories noticed an unusually high number of people dying from the disease, which was almost unheard-of before the 1940s.

The situation in Africa seems to be no better today. Worldwide, an average of 5.9 people per 100,000 will develop oesophageal cancer each year. In East Africa, that figure rises to 9.7 people per 100,000. In Kenya specifically it's 18 in 100,000, while in Malawi it's even higher – 24 in 100,000 – making oesophageal cancer one of the three most common cancers in these countries.

But even after decades of investigation, we still don't really know what's causing these hotspots.

East Africa isn't the only place in the world where this is happening. The

Golestan region of Iran has one of the highest rates anywhere on Earth, and there are pockets of the disease in places as diverse as Henan province in north-central China and southern Brazil, although it's relatively rare in neighbouring Colombia.

Other parts of the world have their own cancer problems: there are strangely high rates of bowel cancer in Slovakia and Denmark, although they have low rates of liver cancer. People in the Czech Republic are more likely to be stricken by kidney or pancreatic cancer than the populations of neighbouring Austria and Poland.

Do these differences lie in inherited genetic variations, or is it something to do with lifestyle? Is there an unknown carcinogen lurking in the environment? Or maybe it's a bit of all three? The wild differences in rates of cancer across the world is a mystery – but a crack team of detectives is on the case.

Leading this team is Mike Stratton, director of the Wellcome Sanger Institute near Cambridge, UK, one of the largest centres in the world for DNA sequencing and analysis. Together with Paul Brennan at the International Agency for Research on Cancer (IARC) in Lyon, France – the World Health Organization's cancer research arm – and other teams in the UK and USA, Stratton has assembled the most impressive detective force in cancer research: a project known as Mutographs of Cancer.

By peering deep inside the DNA of cancer cells, Stratton and his team are hunting for the unique mutational signatures that different cancer-causing agents and processes have left behind.

"I've been interested in the idea that you can detect evidence of the exposures that are causing cancer for 20 or 30 years," Stratton explains. "A mutational signature is simply the pattern of mutations that is left by

a mutational process, and a mutational process can be anything from exposing a cell to ultraviolet light to tobacco smoke to endogenous processes."

The Mutographs team are recruiting 5,000 people across five continents with five different types of cancer, extracting and analysing DNA from thousands of tumours to build up a massive database of mutational signatures – a bit like Interpol's international fingerprint database – so they can try and match causes to cancers around the world.

It's an ambitious £20 million project, funded by Cancer Research UK's Grand Challenge, and is only possible thanks to the international research connections of IARC and the sheer scale of the Sanger Institute's DNA-sequencing pipeline. And its findings have the potential to save many thousands of lives.

At its heart, cancer is a disease of DNA. The human genome contains 20,000 or so genes – the biological instructions that tell our cells when to grow and multiply, what job to do in the body, and even when to die – encoded within long strands of DNA known as chromosomes.

DNA itself is made from four chemical building blocks, or bases, which are strung together in endlessly varied combinations. It's the order of these bases – adenine (A), thymine (T), guanine (G) and cytosine (C) – that conveys the information within a gene, effectively acting like a molecular alphabet spelling out the recipes of life. Any changes to the letters in an important gene – for example, one that drives cell proliferation – might cause a cell to start multiplying out of control.

Further alterations in other vital genes, along with a cellular environment that allows or even encourages unchecked growth, will eventually lead to a tumour. If you can detect the DNA mutations that have led to the development of a person's cancer and work out what caused them, then

you should have the solution to their biological whodunnit. But to do that, you need to be able to read DNA.

In the late 1970s, biochemist Fred Sanger developed a reliable method for reading the sequence of letters in a stretch of DNA, and the institute in Cambridge bears his name as a testament to this game-changing discovery. Sanger's original sequencing technique was time-consuming and cumbersome, allowing scientists to read a couple of hundred bases at best. So rather than looking at all six billion letters of the human genome in search of cancer-causing changes, researchers started by focusing on just one gene, p53, which is faulty in the majority of human cancers.

By the 1990s, Curtis Harris at the US National Cancer Institute and Bert Vogelstein at the Johns Hopkins Oncology Center in Baltimore had managed to show that different types of cancer had their own unique suite of mutations in p53, which were likely to have been caused by different agents, such as the chemicals in tobacco smoke or UV light from the sun.

Stratton – then a young geneticist hunting for mutations in cancers affecting the muscles and other soft tissues – was intrigued by the findings.

"These were very seminal papers which suggested that, yes, mutagens that cause cancer leave their mark on the genome," he recalls. "That had a big impression on me as an opportunity for genomics, but it's one that had to be put away in the locker for 15 years waiting for the technology."

That technology was [next-generation sequencing](#): DNA-reading machines enabling scientists to move from reading hundreds of bases at a time to thousands or even millions. Straight away, Stratton saw the potential for the technology to revolutionise our understanding of the

genetic changes inside individual tumours, setting the Sanger Institute's huge banks of DNA-sequencing machines in motion to read every single letter of DNA in a tumour.

By 2009, he and his team had produced the first whole cancer genome sequences. These were detailed maps showing all the genetic changes and mutations that had occurred within two individual cancers – a melanoma from the skin and a lung tumour.

These choices of cancer types were far from random: decades of epidemiology and lab studies had shown that UV light exposure is likely to be the strongest sole cause of melanoma, while knowledge of the link between tobacco and lung cancer goes back to the 1950s. With such strong lead suspects, Stratton and his team had the best chance of finding clear mutational fingerprints in the genome. But although they were expecting to see the same kinds of mutations across the genomes that Harris and Vogelstein had already picked up in their single-gene studies, they weren't prepared for the sheer scale of genomic vandalism that they uncovered.

"The melanoma had something like 25,000 mutations, which was more than the world had ever seen in one genome," Stratton says. "We could really see the signatures of the exposures that had taken place at an incredibly fine-grained resolution, and we could see all sorts of features and nuances that we hadn't noticed before."

In the same way that a human fingerprint is a mixture of different patterns of ridges, mutational fingerprints are made up of characteristic patterns of DNA changes.

Carcinogenic chemicals cause mutations by physically binding to specific bases and affecting their shape. These alterations throw a molecular spanner in the works, holding up fundamental processes such

as copying DNA or reading genes, so they have to be fixed to keep the cell healthy and functioning properly. For example, benzo(a)pyrene (one of the major carcinogens in tobacco smoke) tends to bind to G bases, as does aflatoxin, a cancer-causing chemical made by certain moulds. But each of these types of damage is repaired in a specific way, leaving a characteristic change in the DNA sequence.

By contrast, UV light leads to mutations by causing neighbouring Cs to become stuck together. When the DNA-copying machinery encounters these fused pairs, it interprets the unusual shape as being a pair of Ts, resulting in a permanent change in the DNA sequence in that position.

"In order to analyse and distinguish between these causes, we have to have a way of classifying the patterns of mutations, a bit like identifying a specific set of fingerprints according to the particular patterns of loops and whorls," Stratton explains.

Initially, Stratton and his team focused on six basic mutational signatures: C to A, C to G, C to T, T to A, T to C and T to G. But there are several different mutational processes that can convert, say, a C to a T, making it difficult to tell what may have been the underlying cause. The researchers then realised that certain mutations tend to appear in the context of certain DNA sequences, as a result of the specific chemical interactions or biological machinery at work.

By expanding out to look at the two bases either side of the mutation – ACA changed to AAA, ACC to AAC, ACG to AAG and so on – Stratton and his team ended up with 96 different subtypes of mutation. Different mutational processes lead to specific patterns across these 96, which pop out of a graphic representation almost as neatly as the ridges and lines of a human fingerprint.

There are also other distinctive changes found in the genomes of cancer

cells – including deletions or insertions of small sections of DNA, characteristic changes to consecutive base pairs, and larger alterations and rearrangements – which can help to further refine the characteristic fingerprint of a particular mutational process.

The melanoma and lung cancer genomes were powerful proof that the fingerprints of specific culprits could be seen in cancers with one major cause. Yet these tumours still contained many mutations that couldn't be explained by UV or tobacco, so what was causing them? And what about cancers without such an obvious single cause? With thousands upon thousands of mutations in a typical tumour, the detective work becomes a lot trickier for cancers with complex, multiple or even completely unknown origins.

By way of analogy, imagine you're a forensic scientist dusting for fingerprints at a murder scene. You might strike it lucky and find a set of perfect prints on a windowpane or door handle that match a known killer in your database. But you're much more likely to uncover a mish-mash of fingerprints belonging to a whole range of folk – from the victim and potential suspects to innocent parties and police investigators – all laid on top of each other on all sorts of surfaces.

Fortunately, a Ph.D. student of Stratton's, Ludmil Alexandrov (now an assistant professor at the University of California, San Diego), came up with a way of solving the problem. He realised that the individual mutational signatures in a tumour can be distinguished from one another using a mathematical method called blind source separation, previously used to separate data from multiple sources, for example splitting out individual vocal and instrumental tracks from a single audio file.

By 2013, the Sanger team had used a version of this technique to extract 20 distinct mutational fingerprints from nearly 5 million mutations in more than 7,000 tumours, covering 30 of the most common forms of

cancer. Some fingerprints turned up in every single tumour, while others were specific to just a handful of cancer types. All of the cancers had at least two different fingerprints, while some had at least six.

That number rose in 2015 to at least [30 unique mutational fingerprints](#), each caused by a different agent. Then in 2018 an [even larger analysis](#) of nearly 85 million mutations in around 25,000 cancers raised the number of fingerprints to around 65, although probably only around 50 of these are truly unique.

Some of them come from things we already know can significantly increase the risk of cancer – the usual suspects like tobacco or PAHs (polycyclic aromatic hydrocarbons, released when certain materials burn). Some previously suspected carcinogens have also been confirmed as dangers, such as aristolochic acid, a chemical produced by plants that were commonly used in herbal supplements in Taiwan and elsewhere. Other fingerprints are signs of an inside job, resulting from the fundamental processes of life inside our cells, including DNA copying and repair.

But the causes of around half these fingerprints remain a mystery, left in the genomes of cancer cells by culprits that are still at large.

The endoscopy suite in Moi Teaching and Referral Hospital in Eldoret, western Kenya, is a busy place. Oesophageal cancer is one of the most common tumours around here, and every day a seemingly endless stream of patients arrive in search of relief from the bulging blockages in their gullets. Most have not eaten properly for weeks, making them incredibly thin and frail. Some have come hundreds of kilometres from the outlying rural areas, spending precious money on pricey hired transport. All of them are desperate for help, and most are going to die within the next year.

I watch as little pink globs of cancer tissue are carefully popped into plastic pots and sent off to a freezer in a building on the other side of the hospital, waiting to be shipped to IARC. The team there will purify the precious DNA from each sample, then send it to the Sanger Institute to be sequenced and analysed for any tell-tale mutational signatures that might explain what is causing all these cancers.

I'm here in Eldoret to meet Diana Menya, a Kenyan epidemiologist who has worked at the Moi for many years. She's generous and good-humoured, with an infectious energy. We feel like old friends after bouncing along in a minibus for just two days – including an impromptu side trip to watch teenagers fling themselves off a cliff into a crocodile-infested pool – and she is full of laughter and stories about the region and its inhabitants. It was this curiosity and passion for her local area that first alerted her to the unusually high rates of cancer in the region.

"Some years ago, I noticed that there were quite a number of patients presenting with difficulty swallowing, and when the diagnosis was finally done it was squamous cell oesophageal cancer," Menya explains. "We were seeing more and more patients coming into the hospital, and I was wondering: What is this? What is happening here? Something needs to be done."

Her solution, collaborating with researchers at IARC, was to set up [ESCAPE](#) (Oesophageal Squamous Cell Carcinoma African Prevention Research), a case–control study recruiting people who have oesophageal cancer and people who don't, to compare their environments and lifestyles. By working with Menya, the Mutographs of Cancer team have been able to get hold of tumour and blood samples for DNA analysis and then match them to the information gathered by the ESCAPE team about the environmental or lifestyle factors that might be at work in the region.

Menya's study has already found that tobacco and alcohol are two of the factors likely to be responsible for the surfeit of oesophageal cancers in western Kenya – not entirely surprising, given that previous epidemiological studies have linked them to squamous cell carcinoma. But although these two culprits may account for a lot of cases, they certainly can't explain them all.

Heading out into the rural community outside Iten feels like walking into a world of carcinogens. The farmland may be lush and fertile here, but it is also awash with pesticides and fertilisers that can leach into the (usually unfiltered) water supply. There are sprouting fields of collard greens that are cooked into a dish known locally as sukuma wiki ('lasts a week') that's particularly high in nitrates, which may then be converted into carcinogenic nitrosamines in the body.

We visit Emily in her kitchen – a single unventilated room coated with a thick layer of soot that hangs down in some places like stalactites. The smoke from the open fire is overpowering, and it's impossible to stay in there for more than a few seconds without dashing out for air. And where there's smoke, there's PAHs, released from burning fuels such as wood, maize cobs and cow dung. Women, young girls and children are particularly exposed as they spend so much time in the kitchen, often sleeping in there at night to keep warm and safe.

Maize is a common food and fuel source in this area, and the cobs and kernels are often treated with fungicide to prevent the growth of toxic pink mould (which may itself cause cancer). Burning these chemicals along with the cobs may release further carcinogens into the unventilated atmosphere.

Then there are personal habits. It's common in East Africa to drink extremely hot tea, sipped at mouth-scalding temperatures of up to 70°C (something I discovered the hard way). Very hot drinks have already

been linked with oesophageal cancer in Iran and parts of South America.

Poor dental hygiene could be another factor. A study in China has shown that the fewer teeth a person has, the greater their risk of oesophageal cancer, perhaps as a result of toxic bacterial chemicals leaching into the saliva from infected gums.

Although it's easy to suspect all these things (and more) as being behind the high oesophageal cancer rates in Kenya, we don't yet have enough data to link most of these risk factors with fingerprints left in the cancer genome. That means we don't yet know which of them are the most dangerous, or how they might act together to cause disease. In the future, we should be able to match more carcinogens with their fingerprints by combining the Mutographs approach with diligent epidemiological studies like ESCCAPE. But this isn't the only way.

Rather than studying DNA extracted from tumours to look for mutational signatures, David Phillips, a professor of environmental carcinogenesis at King's College London, is coming at the problem from the opposite direction. As part of the Mutographs project, he and his team are treating laboratory-grown cells with DNA-damaging agents, then sequencing their DNA to see what mutational signatures have been left.

"By looking in a systematic way in human tumours and comparing them with mutational signatures in experimental systems that are caused by things we think or know are carcinogenic to humans, we can match the two and say, 'Aha! Here is evidence that this particular chemical is involved,'" Phillips explains. "We're working our way independently through things that we suspect or know cause human cancers and seeing what signatures we can generate from those."

So far Phillips and his team have tested 80 suspected causes of DNA

damage, of which around half produce distinctive fingerprints in the genome. Some are known human carcinogens, such as UV light and aristolochic acid, which produce patterns of damage that would be expected based on their properties. Whenever those turn up in the Mutographs tumour samples, it's a fairly safe bet that the relevant agent is involved somewhere.

But he's found other chemicals that leave mutational signatures in lab-grown cells which haven't yet been detected in human tumours. Maybe these molecules are genuinely carcinogenic but it's rare that people get enough exposure for it to show up in their [cancer cells](#), or maybe they're truly innocent and can be ruled out of Stratton's investigations.

It's a bit like catching someone speeding in a car, taking their fingerprints and putting them in a database: maybe they will never go on to commit a more serious crime, but if they do, then the police have a much better chance of catching them.

Right now, we're still in the opening pages of this detective story. A year after the Mutographs project began, a full team has been assembled, the tools are in place, and the researchers are starting to gather and analyse fingerprints from cancers all over the world. The sheer scale of the project is staggering.

"As we refined the sophistication of the approach and the mutation classification, the algorithms and the sequencing, it has become clear that this was a big challenge that would require coordinated investment and organisation," says Stratton. "We have to collect five to ten thousand tumour samples and normal blood, we have to quality-control the DNA sequencing and do the data management and statistics – it's a combination of large-scale epidemiology and large-scale genomics that haven't been married together in this way before."

In June 2018, researchers from the Mutographs team gathered at the Sanger Institute to share preliminary data from the first handful of cancer genomes to make it through the project's pipeline. Intriguingly, the first few oesophageal tumours from Kenya don't appear to have any signatures from PAHs, potentially putting smoky rural kitchens like Emily's in the clear, although it's important to stress that only a small fraction of cases have been analysed so far.

Curiously, all the oesophageal cancers do have signs of damage caused by APOBECs – DNA-altering proteins in our cells that are thought to be activated in response to viral infections. There's a lot of scientific hand-waving about the role that these internal mutators might play in cancer, and even less is known about what triggers their activity in the absence of viruses, but the discovery of their fingerprints at the biological scene of the crime is intriguing.

While it's too early to nail down any suspects for sure, in a plot twist worthy of Agatha Christie's *Murder on the Orient Express* – SPOILER ALERT – it's becoming clear from the cancer genomes that we're not dealing with individual baddies but a gang of miscreants, each of whom administers a potentially fatal blow to the genome. Each causes mayhem in its own way, but they can combine to bring about catastrophe.

The flipside is that it's almost impossible to pinpoint one specific cause of a given tumour or say exactly what caused it. A cell may be riddled with mutations, accumulated from all sorts of processes over a lifetime, but if none of them hit the vital genes or control switches responsible for growth or death, then it will remain healthy. And because every cancer genome is shot through with many thousands of mutations, it's impossible to say which culprit delivered the coup de grâce. But we should be able to build up a much better picture of the contributions of different factors – be they biological or environmental – to each individual person's disease.

This research is coming too late to help Emily's parents, lying in their hillside graves. But Diana Menya is hopeful that her work will save lives in the future.

"I think it's a combination of environmental factors and behavioural factors. Something to do with what people do, what they eat, how they live and their habits, such as tobacco consumption and alcohol consumption. It could also be something like indoor air pollution," she says, calling to mind women like Emily and her smoke-filled kitchen. She's also hopeful the project will pick apart the roles of nature and nurture in causing Kenya's [oesophageal cancer](#) epidemic. "We've got anecdotal information that it runs in families in our study site, but is it genetics or is it a shared environment? I'm hoping the Mutographs study can answer that question."

This is real prevention research – finding out what is increasing the risk of cancer at a fundamental level, then using the knowledge to make lasting life-saving changes in public health. But this takes time and it needs political will. Menya has already seen this approach bear fruit in her own work running a campaign to support the extermination of guinea worm across the country.

"To prevent a disease, we have to take action. We have eradicated many conditions, and especially infectious diseases, so can we not eradicate non-communicable diseases like cancer?" Menya says. "This is going to sound quite ambitious, but for me the best endgame will be a cancer-free Kenya by preventing cancers from occurring in the first place – catching it before it catches the patients."

This [article](#) first appeared on [Mosaic](#) and is republished here under a Creative Commons licence.

Provided by Mosaic

Citation: The DNA detectives hunting the causes of cancer (2018, September 25) retrieved 19 April 2024 from <https://medicalxpress.com/news/2018-09-dna-cancer.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.