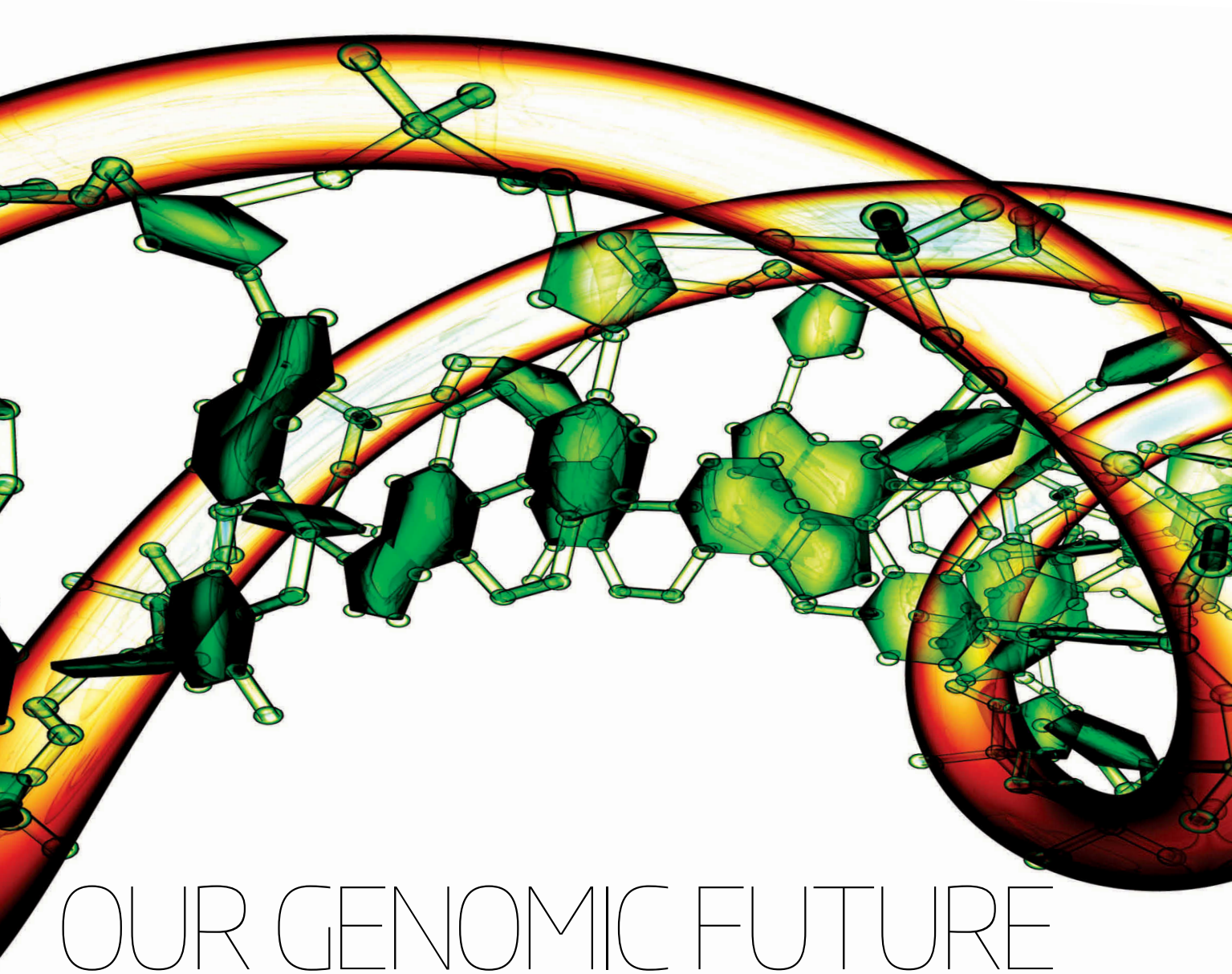

PASIEKA/SCIENCE PHOTO LIBRARY

OUR GENOMIC FUTURE

Genome sequencing is delivering a medical revolution that should one day come to us all, as Peter Aldhous discovers

FOR the Yuska family, the future of medicine is here. Thanks to genome sequencing, parents Danielle and Erik have a name for the mysterious condition that they feared would take the life of their 7-year-old daughter, Lillian, and they have an idea of what her outlook might be.

Born prematurely, Lillian struggled to feed, suffered from chronic vomiting and diarrhoea, and succumbed to repeated infections. After shuttling for years from specialist to specialist, the Yuskas now know that Lillian has trichohepatoenteric syndrome-2, caused by a mutation in a gene called *SKIV2L*, which disrupts both gut function and immunity. Just six other children worldwide are known to have the condition. "It's real rare, and little is known about it," says Danielle.

Families afflicted by rare genetic disorders will be the first beneficiaries of genome sequencing. But everyone else will benefit, too, and more than you might imagine. It turns out that conditions any of us may face – such as heart disease and diabetes – may also be triggered by similarly rare mutations.

> Most common diseases are really a collection of rare ones with the same end point

"Most common diseases are merely a collection of rare diseases that happen to have the same end point," says David Dimmock, a paediatric geneticist at the Medical College of Wisconsin (MCW) in Milwaukee. In other words, common conditions are actually rare disorders in disguise, which means that the same methods that helped the Yuskas promise to deliver personalised medicine for us all.

To witness this emerging revolution, I visited the pioneering clinic that delivered Lillian's diagnosis, run by

MCW and the Children's Hospital of Wisconsin, and began an exploration of my own DNA that hints what might be possible – and the challenges ahead.

Over the past three years, the Milwaukee clinic has sequenced the genomes of around 40 patients. This has boosted its rate of diagnosis of rare genetic disorders from around 10 per cent to 27 per cent (*Science Translational Medicine*, doi.org/nh8).

The clinic got off to a spectacular start with its first patient. At 4 years of age, Nic Volker's inflammatory bowel disease was so severe that food in his gut would leak into his abdomen. The Milwaukee team zeroed in on a mutation in *XIAP*, a gene that, when damaged, can cause a leukaemia-like disorder.

To avoid that outcome, and suspecting that the defect might also be causing Nic's immune system to attack his bowel, his doctors recommended a bone-marrow transplant (*Genetics in Medicine*, doi.org/dgf8qf). It worked, and while it was too late to save Nic's colon, he is today leading a healthy life.

Not every diagnosis has led to a similar breakthrough. But for families like the Yuskas, ending the diagnostic roller coaster is in itself a huge step forward. For a bleak couple of weeks, the leading contender for Lillian's condition was Fanconi anaemia, which typically kills people in their 20s. "After that, we decided not to research the things she was being tested for," Danielle says.

Lillian's actual diagnosis holds more hope: while her health problems won't go away, there's no reason to conclude that her life will be cut drastically short. "Now we need to find out how best to care for her," says Danielle.

Even when the news is the worst possible, families have been grateful to know what's wrong. Dimmock recalls telling the mother of an infant with severe liver disease that the mutations responsible would quickly lead to fatal neurodegeneration. The mother cried, then turned to Dimmock and said: "Thank you. That really helps." The diagnosis meant that her daughter could be spared a liver transplant that would have proved useless.

Although I'm in good health, I asked the Milwaukee clinic to analyse my "exome". Essentially, the exome is the part of the genome that is "read" to make proteins. While it accounts for only about 1 per cent of the total DNA (the rest regulates gene activity, or is "junk", without a known function), it is thought to host around 85 per cent of the genetic variants that cause disease. I'd previously had my genome scanned by companies that use DNA "chips" to analyse about 1 million common variants, but an exome contains far more information.

CHILDREN'S HOSPITAL OF WISCONSIN

Only six other children in the world are known to have the same condition as Lillian Yuska

Geneticists had hoped that heritability of common conditions like heart disease would be explained in large part by a few relatively common genetic variants. It hasn't turned out that way. Instead, the genetic roots of common disorders seem to be spread across a plethora of rare mutations – some of which may even be unique to their carriers.

This transforms the process of genetic discovery. To find common disease-causing variants, you might recruit a thousand people with the condition in question, plus a similar number of controls, and look for variants that seem to be inherited with the disease. This is how geneticists know, for instance, that my single copy of the variant ε4 in the *APOE4* gene gives me – together with about 1 in 5 people of European ancestry – about twice the typical risk of ›

WHAT I FOUND IN MY DNA

To peek into the future of personalised medicine, I had the protein-coding regions of my genome – the exome – sequenced and analysed. The vast amount of data highlights the challenges that have to be overcome before this technique becomes routine

My DNA, in numbers
- One exome
- About 30 million DNA bases (of 3 billion in the entire genome)
- 153,984 differences from the reference human genome
- 61,454 variants passed quality-control procedures intended to exclude errors
- 8366 would alter the protein made from the gene

What does this mean for my health? Here are a few of the more interesting variants

LOCATION: Chromosome 11
GENE: *PYGM*
A variant in this gene indicates that I am a carrier of McArdle disease, which makes those who have it intolerant to exercise. It prevents glycogen, used to store energy in our muscles, being converted into glucose. I don't have this condition, as you must inherit two copies, so can't use it as an excuse to be lazy

LOCATION: Chromosome 6
GENE: *HFE*
Along with about 1 in 10 Europeans, I am a carrier of haemochromatosis. There is also a small chance, less than 1 in 50, that I could develop the condition, in which the body accumulates iron, leading to liver damage and other problems. I could watch for this by testing my blood for an iron-containing protein called ferritin. If my levels are high, giving blood from time to time would fix the problem

LOCATION: Chromosome 19
GENE: *APOE*
I have one copy of the ε4 variant, which roughly doubles my risk of developing Alzheimer's disease

LOCATION: Chromosome 10
GENE: *CYP2C19*
If I ever need to be treated with an anti-clotting drug, clopidogrel may not be a good choice thanks to a variant that disrupts conversion to its active form

LOCATION: X Chromosome
GENE: *ABCD1*
Thankfully, the variants of this gene that appear in my unfiltered data are glitches. If they weren't, I would probably have died of X-linked adrenoleukodystrophy, which causes neurodegeneration. The glitches were caused by a "pseudogene", a stretch of mutated DNA that was once a gene but which no longer functions, being mapped onto *ABCD1* in error. It was flagged as failing quality control by 23andMe, the company that sequenced my DNA

Source: Medical College of Wisconsin, plus analyses using Omicia Opal and Ingenuity Variant Analysis

developing Alzheimer's disease.

But if a thousand people with the same external symptoms have hundreds of different mutations causing them, finding the gene variants responsible becomes a needle-in-a-haystack problem.

That might sound hopeless, given that a typical genome contains several million departures from the reference human DNA sequence. But you can winnow down the list of candidates using algorithms that predict whether a variant is likely to disrupt a gene's function. A premature "stop" instruction, for instance, which cuts a protein short, is an obvious candidate for further scrutiny.

These algorithms work in various ways. Some flag up mutations if they occur in DNA sequences that differ little from species to species. The thinking goes that if evolution has preserved a gene's sequence, then any change in it is likely to be bad.

Other algorithms are more sophisticated, for instance assessing how gene variants might change a protein's shape, which might affect how it works.

GARY PORTER/THEMILWAUKEEJOURNALSENTINEL

Thumbs up for Nic Volker's diagnosis

The Milwaukee clinic's bioinformatics team, headed by Liz Worthey, runs five of these algorithms. To get a sense of what they do, I ran three on my exome, producing a list of almost 1700 potentially damaging variants. That might sound worrying, but it's actually normal. "Healthy people tend to have healthy genomes," I was reassured by David Bick, a clinical geneticist on the Milwaukee team, who talked me through the results in the presence of a genetic counsellor.

A mutation may be damaging to a protein without causing disease. This is because many of our metabolic pathways have a degree of redundancy, which means that when individual proteins "break", there are often others available to take the strain.

What's more, some of the variants highlighted by the prediction algorithms will be sequencing errors. Quality control "proof-reading" of the sequence is good at spotting obvious errors. However, genome sequencing technologies, while rapid and cheap, currently lack the accuracy of the old-school method known as Sanger sequencing. This is why the Milwaukee team confirms every variant it suspects of causing disease using the Sanger method.

While sequencing methods and the tools used to sift for damaging mutations have advanced far enough to spot single variants causing rare diseases, the signal-to-noise problem becomes more challenging when the symptoms are common. And as I can vouch, it's possible to spend days exploring what has been dubbed the "incidentalome" – variants which look as if they might be damaging, but which are probably of little significance. As our knowledge of rare variants expands, though, and prediction algorithms become more sophisticated, the genome's dark matter should come into view.

Even at this early stage, there are valuable nuggets to be mined from a typical genome – as I found (see "What I found in my DNA", page 11). The most useful information surrounds carrier status for genetic disorders and clues about how commonly used drugs are likely to work.

The full benefits of the genomic revolution will come when doctors are routinely able to pick out the rare variants triggering certain symptoms, and prescribe drugs that have been specifically designed to correct the biochemical pathways concerned.

"Hundreds of different gene mutations could be causing the same external symptoms"

This may take several decades. Not only must geneticists get better at deciphering the significance of specific variants; drug companies will also need to shift their focus from developing blockbuster drugs sold to millions to supplying the niche markets of personalised medicine.

Thanks to the trail blazed by the Milwaukee clinic and a handful of others, that journey has begun. As Howard Jacob, who heads MCW's Human and Molecular Genetics Center, told me: "Your genome is going to get more and more meaningful." ■