

GENOME DIAGNOSTICS IN THE FAST LANE

Streamlined workflows for DNA and RNA sequencing are helping clinicians to deliver prompt, targeted care to people in days – or even hours. **By Michael Eisenstein**

Sometimes good news arrives too late. About a decade ago, clinical geneticist Zornitza Stark and her colleagues at the Murdoch Children's Research Institute in Melbourne, Australia, set out to learn how genome sequencing might improve the care of young children with suspected hereditary disorders. The researchers sequenced the protein-coding regions in the genomes of 80 infants, looking for genetic changes to explain their conditions. For one nine-month-old boy with severe neurological symptoms, Stark's team homed in on a mutation that resulted in a crucial vitamin B1 deficiency, which could be readily treated with supplements for just under Aus\$150 per year (US\$136 at the time).

Unfortunately, it took six months from when the sample was taken to get that answer. Treatment with vitamin B1 safeguarded the boy's survival, but his condition had deteriorated in that time, leaving him with a lifelong disability. "Even with a two-, three-week turnaround time, that's clearly still not fast enough to intervene with a lot of these conditions," says Stark. "I was very keen to make it even faster."

Today, the method is indeed much faster. Stark and others around the world have repeatedly demonstrated the feasibility and clinical benefit of rapid sequencing and interpretation pipelines for delivering timely, personalized interventions for previously enigmatic disorders. These workflows can generally deliver diagnoses in less than a week, but the record for a genome-sequencing-derived diagnosis, held by genomics researcher Euan Ashley at Stanford University in California and his colleagues, is just over seven hours¹. And when an ultra-rapid response is needed in oncology, workflows are even helping clinicians to classify tumours and plan treatment in under an hour.

Sequencing-based diagnostics can be a hard sell to some clinicians, particularly those inclined to take a conservative approach – and until a decade or so ago, genome analysis was slow, expensive and unlikely to deliver actionable results. But the sceptics are coming around. "When I talk to clinical collaborators, they always start with, 'Ah, we don't need it,'" says Jeroen de Ridder, a bioinformatician at University Medical Center Utrecht in the Netherlands who developed a sequencing-based approach for guiding brain-tumour



Ultra-rapid sequencing of human genomes can identify rare diseases and speed up treatment.

care within 90 minutes of biopsy collection². "But then, when you enable it and you make it possible, suddenly they start to see opportunities that they previously couldn't see."

The technology is advancing so quickly that even early adopters can find it hard to keep up. "Every 18 months or so, we have a technological revolution – not evolution, revolution," says Stephen Kingsmore, a rapid-sequencing pioneer at the Rady Children's Institute for Genomic Medicine in San Diego, California. "We're in the middle of that right now."

Breaking through the bottlenecks

Genetic diagnostic assays have been guiding clinicians for decades. Early tests focused on genes with known ties to disease, such as the breast cancer risk-factor gene *BRCA1*. But as sequencing costs have fallen, whole-genome testing has become increasingly popular for diagnosing cancers or rare genetic disorders. The wheels of this process grind slowly, however. "Nothing about clinical genetic testing was set up with speed in mind," says Ashley.

Sequencing billions of nucleotides with sufficient 'redundancy' to ensure high-accuracy results takes time – and scrutinizing the results to identify potentially relevant sequence aberrations takes even longer. Even a more targeted 'exome' analysis that focuses on the 1–2% of the genome containing protein-coding genes takes months to deliver answers.

Kingsmore describes his first attempt to accelerate the process as "serendipitous". In 2012, sequencing firm Illumina in San Diego, reached out to Kingsmore's institution at the time – Children's Mercy Hospital in Kansas City, Missouri – to test its new 'rapid run mode'. Illumina was, and is, the market leader for short-read sequencing, a process that produces billions of 100–200-nucleotide 'reads' of DNA sequence, which can then be computationally reassembled into whole genomes or exomes. This mode combines advances in sequencing chemistry and analytics to streamline and reduce the sequencing step from about 10 days to one. Kingsmore and his colleagues built a diagnostic workflow

Work / Technology & tools

around the new mode that allowed them to identify probable disease-causing mutations in critically ill newborn babies within 50 hours of taking a blood sample³.

“Three of the first four cases that we decoded genomes for had findings, and it was like a bomb went off,” says Kingsmore. Since then, he has continued to refine the process, publishing dozens of studies and coordinating a trial called Baby Bear that demonstrated the feasibility of implementing this diagnostic workflow with nearly 200 infants at 5 California hospitals between 2018 and 2020. He estimates that thousands of children in the United States now receive this service annually.

Other groups have developed similar programmes. Stark’s team is coordinating the Acute Care Genomics programme for babies and children across all Australian hospitals, while clinical geneticist Ahmad Abou Tayoun at the Al Jalila Children’s Specialty Hospital in Dubai, United Arab Emirates, is leading the Little Falcon programme, which will carry out rapid sequencing for 200 Middle Eastern families with critically ill infants. In the United Kingdom, clinical geneticist Emma Baple at the University of Exeter is coordinating a National Health Service programme to provide rapid sequencing for 1,200 families in England and other parts of the United Kingdom each year.

Built for speed

To achieve these accelerated workflows, clinical genomics teams have identified multiple opportunities to streamline the process. For one, they’ve mostly eliminated exome sequencing, which requires extra sample preparation steps and sacrifices valuable information. “We pick up about 3–4% of diagnoses that wouldn’t have been detected by exome sequencing since we moved to whole-genome sequencing,” says Baple. This includes crucial mutations that reside outside coding regions that affect gene expression and regulation.

Sequencing instruments have also become much faster. Many labs use Illumina’s state-of-the-art NovaSeq X Plus, which can churn out trillions of nucleotides of sequence data per day. But this can be costly – the economic advantages of such high-throughput systems come when many samples are processed at once. But most rapid-sequencing samples represent urgent individual cases, and so doing the sequencing costs as much as it does for dozens of genomes in a more efficient workflow.

Whatever the workflow, these rapid-sequencing programmes achieve remarkably consistent success, typically delivering a confident diagnosis in 40–50% of cases, many of which can be addressed medically. Interpretation can still be challenging, but this too is becoming easier thanks to the millions of clinically annotated genomes and exomes collected through initiatives such as the UK Biobank programme and the

Genome Aggregation Database.

Also helping is an ever-advancing computational toolbox. Preliminary genome processing and analysis can now be performed using cloud-based resources, sparing researchers the need for costly high-performance computers. And software packages have evolved to automate crucial analysis steps. Both Kingsmore and Abou Tayoun are using Illumina’s DRAGEN toolkit to quickly converge on relevant sequence and structural variants, for instance. There are also useful resources such as GeneMatcher, an online service that helps geneticists and clinicians around the world to exchange insights about disease-related mutations. Within minutes of using GeneMatcher to

“Three of the first four cases that we decoded genomes for had findings, and it was like a bomb went off.”

help diagnose one young patient, Baple’s team was connected to reports of similar mutations in other children with a heart condition called dilated cardiomyopathy. “That record of dilated cardiomyopathy meant that child could potentially get onto the transplant list when they were big enough,” she says.

And then there is the promise of artificial intelligence (AI). Abou Tayoun estimates that geneticists and other specialists take about an hour to manually review each genome before the first diagnosis is delivered. But AI capabilities are maturing and it won’t be long before

the technology has a major impact, Kingsmore predicts. AI systems “do miss things, and they do make mistakes, but you can now get a 10-minute interpretation on the genome,” he says. In the meantime, “we still have a lot of people power involved”.

Taking the long view

Illumina remains the backbone of almost all clinical-sequencing programmes worldwide, but it faces a host of competitors. Although still too new to have meaningful clinical track records, these platforms could open up rapid sequencing to users who cannot afford the high cost of single-sample analysis on the NovaSeq X Plus, which has a price tag of more than \$1 million.

Kingsmore is especially excited about the AVITI instrument from Element Biosciences in San Diego. The instrument is designed for medium throughput and to deliver per-base accuracy that the company says matches or exceeds that of Illumina. “You buy three instruments – and the instruments are relatively inexpensive, about \$250,000 each – and you’ll get \$200 genomes,” he says. By comparison, sequencing a full genome cost roughly \$10,000 just a decade ago. “That is going to be quite transformative.”

But perhaps the greatest excitement centres on ‘long-read’ technologies, particularly those developed by UK firm Oxford Nanopore Technologies (ONT). Long reads are valuable for documenting gene duplications, deletions and other structural changes that are challenging to capture with short-read methods. Importantly, this technology can also be used to



Infants with unknown conditions can benefit from fast diagnoses enabled by rapid sequencing.

detect other 'omics' indicators when an ultra-rapid response is required, such as for cancer (see 'A matter of minutes').

ONT instruments generate sequence data from DNA fragments that span thousands of bases by monitoring fluctuations in electrical current as the DNA travels through a tiny pore. Sequencing runs can be completed in just a few hours and can even be analysed as data flow into the computer. "It's a massive leap in terms of the sequencing time," says Stark, who hopes to start evaluating ONT in the next year or so. And, she adds, in terms of flexibility, especially when it comes to the ONT MinION instrument, which is roughly the size of a mobile phone. "Those sequencers are often portable, so what would be really exciting is being able to bring the sequencing closer to the bedside."

Ashley's team used ONT's high-throughput PromethION bench-top system to achieve its record-setting genomic diagnoses. By dedicating an entire sequencing run to one sample, the team could reduce total sequencing time to just a few hours. But several bottlenecks remained, and even a high-powered computer with a state-of-the-art graphic processing unit required 18 hours of analysis time per sample.

"We really looked at every part of the pipeline and said: is there a way we can do this faster?" Ashley recalls. This meant stripping down sample preparation to bare essentials, and transitioning to a cloud-based data-processing pipeline to run algorithms such as DeepVariant — a deep-learning tool developed by Google for rapidly identifying and classifying potential disease-related genetic variations⁴. In a best-case scenario, the time required for sequence interpretation "could be as little as ten minutes," says Ashley, although more challenging cases could take longer if thorough manual evaluation is needed. His team is now looking at scaling up and accelerating this process, and several other groups are beginning exploratory work with ONT. "I think we're getting an extra 10–15% additional diagnostic findings, mostly structural and methylation variants," says Abou Tayoun of his group's experiences.

Rebooting the system

But when it comes to clinical diagnostics, technology is only half the battle — there are also the human and logistical elements to deal with. "The bottleneck is really to get the doctors to be familiar with the fact that it's available and when to order it," says Abou Tayoun. "And of course, making sure that when they get the results, they can act on them."

This entails working out a variety of other details, including the optimal balance between speed and cost; and whether delivering a diagnosis in less than one day meaningfully improves outcomes relative to three or five days. For critically ill newborn babies with severe respiratory, heart or liver problems,

A matter of minutes

Ultra-fast sequencing could guide efficient cancer management, too.

Although mainly a tool for DNA sequencing, platforms developed by UK firm Oxford Nanopore Technologies (ONT) can extract other types of molecular information, including gene-expression patterns and epigenetic modifications to chromosomal DNA. Such data can help clinicians to accelerate cancer diagnosis, allowing crucial treatment decisions in hours rather than weeks.

Paediatric brain tumours, for instance, can manifest in multiple forms, some of which are more aggressive and treatment-resistant than others. "The question for the surgeon is, how aggressively am I going to operate?" says Jeroen de Ridder, a bioinformatician at University Medical Center Utrecht, the Netherlands. "Am I going to go all-in and take some risk, or can I be more conservative?" DNA sequence data can help to categorize such tumours, but epigenomic markers — especially DNA methylation patterns — are much more informative, says Stephen Kingsmore, a rapid-sequencing pioneer at the Rady Children's Institute for Genomic Medicine in San Diego, California.

A team co-led by de Ridder developed a deep-learning algorithm called Sturgeon, which the researchers trained on millions of simulated data sets derived from patient

the urgency is high, but there are considerations beyond the sequencing itself. "It's all very well to speed up the test to be done in under 10 hours or whatever, but if it takes 3 weeks for that child to even access the test, you're not really going to get the full benefit," says Stark.

Equitable access remains a formidable problem. Moving data analysis to the cloud lifts the computational burden from hospitals. But without sequencing equipment and skilled users on site, resource-limited health-care facilities might have to send specimens to centralized labs, introducing delays and logistical hurdles. This could give the affordable and portable ONT instruments an edge, says Ashley — particularly in rural communities or low- and middle-income countries with limited lab resources.

As rapid-sequencing workflows mature, these issues should resolve. And the community of people who could benefit is huge. Kingsmore is now applying his team's rapid genome sequencing to screen tens of thousands of

DNA methylation profiles. Last year, the team reported that Sturgeon accurately classified 45 out of 50 tumour samples after just 40 minutes of sequencing on an ONT MinION device². That's fast enough for clinicians to make treatment decisions while patients are still under anaesthesia for the biopsy, and to proceed directly to aggressive surgery if needed. The team has also tested the approach in a true intraoperative setting. A similar approach could prove useful for other heterogeneous tumour types, including soft-tissue sarcomas, de Ridder says.

Scientists are also developing sequencing-based 'liquid biopsies', in which body fluids such as blood and urine are tested to rapidly detect, monitor and assess treatment response in cancer. A team led by computational biologist Martin Smith at the University of New South Wales in Sydney, Australia, has shown that RNA sequencing with the MinION could shorten the time for leukaemia profiling of blood samples from days to minutes⁵. "Five minutes was enough for most of the samples to get above 90% classification — that's only 30,000 reads," says Smith. The algorithm could distinguish 11 subtypes of acute lymphoblastic leukaemia on a standard laptop, and Smith says that a similar approach could be applied to other time-critical conditions, including sepsis, infections and autoimmune-disease flare-ups.

newborns for nearly 400 genetic disorders as part of a study known as BeginNGS. This project, he says, is partly a testing ground for new technologies to strengthen the process, including AI-guided tools that currently deliver roughly 90% of the diagnostic acuity of human experts. But more generally, it's about paving the way to a future in which anybody can benefit from fast, accurate sequencing.

"We need something we can do on millions of samples per year for less than \$1,000, and maybe less than \$500," says Kingsmore. "We have to think about automation on a completely different scale."

Michael Eisenstein is a science writer in Philadelphia, Pennsylvania.

1. Gorzynski, J. E. *et al.* *N. Engl. J. Med.* **386**, 700–702 (2022).
2. Vermeulen, C. *et al.* *Nature* **622**, 842–849 (2023).
3. Saunders, C. J. *et al.* *Sci. Transl. Med.* **4**, 154ra135 (2012).
4. Goenka, S. D. *et al.* *Nature Biotechnol.* **40**, 1035–1041 (2022).
5. Sagniez, M. *et al.* Preprint at medRxiv <https://doi.org/10.1101/2022.06.22.22276550> (2022).