Genomics for the Smart Nation

By Swaine Chen and Andreas Wilm

Science and imagination have an intimate relationship. Stories and movies both draw on and also drive forward real research. In transportation, for example, we imagined flight and space travel – and both became reality. In space, we dream of faster-than-light travel and physical teleportation (instant travel!). If we can't physically teleport ourselves, the next best thing is instant communication – and we're lucky to live in a time where voice, video, and even holograms are routine technologies. Interestingly, everyone today has the ability to broadcast live video – yet we still get stuck in traffic, with many taking longer to commute to work than they did just a few years ago. What differentiates the progress of digital communication from that of physical transportation?

The Revolutionary Progress of Computing

The difference, we think, is that communications has leveraged on the power of computers. Staggering increases in computing power (and decreases in cost) have enabled digital devices to pervade modern life. Anything that becomes digital, anything that can be reduced to a transaction of information, has been on an exponentially accelerating curve of progress.

One great formula for realising our imagined future, therefore, is to ride on computers. This is one of the foundational drivers for Singapore's Smart Nation initiative. Ubiquitous computing is possible because it has become so cheap; miniature sensors can be digitally connected and leverage communications networks. The benefits for better management and higher efficiency are clear for daily life – shopping, commuting, and socialising.

Healthcare, particularly through genomics, is also ripe for transformation through computing and big data. We work at the Genome Institute of Singapore (GIS), a national flagship programme for the study of genomes and genomics. A genome refers to the complete set of DNA that is present within an organism, including within every cell of our bodies. It is the "blueprint" of life. Our DNA makes us human; lions have different DNA that makes them lions.

Computers work on binary bits – 0 and 1. DNA works on a four-base code – G, A, T, and C. Four different bases are equivalent to two bits, making DNA a natural domain for computation. And DNA doesn't just differentiate humans from lions; DNA can encode differences in health risk. An individual's DNA determines whether they have a disease like cystic fibrosis, muscular dystrophy, or sickle cell anaemia. We can use DNA to diagnose Down's syndrome; assess breast cancer risk; and to determine who might get side effects from a drug. GIS is working to further decode what our DNA means, leading to an exploding ability to use DNA to diagnose, treat, and even predict disease. It is therefore on the forefront of converting DNA analysis to a computational task – the secret formula for transformative progress mentioned above, which has converted the imaginations of yesterday into today's realities in global communications.

In our own work, DNA analysis by computers has revolutionised how we investigate infectious outbreaks. In two recent examples in Singapore, genomics helped to prove that the Group B Streptococcus outbreak in 2015 arising from the consumption of raw fish was caused by a unique strain of bacteria present only in South East Asia. Similarly, our colleagues at A*STAR used genomics to show that the 2016 Zika outbreak in Singapore was not caused by the viruses circulating in the Americas that were associated with birth defects.

An Even Bigger Revolution

We use the word "exploding" for genomics because, amazingly, genomics in recent times has been progressing even faster than computing. The pace of genomics is breathtaking to the point that we struggle to even describe it. It's as if we leapt across 50 years of progress – 1960s mainframe computers to 2010s handheld smartphones – in less than 10 years.

The advent of cloud computing has dovetailed perfectly in time and scale with genomics. At GIS, we are experts in DNA analysis, not building and maintaining computers. Cloud computing – which uses remote servers hosted on the internet to store, manage, and process data – allows us to access a global computational infrastructure which we couldn't (and shouldn't) reproduce ourselves. In doing so, we focus more on our core expertise of DNA and genome analysis while saving resources and time. Together, this may mean faster detection of outbreaks, leading to fewer people getting sick and an overall safer food supply.

The trend of ascension into the cloud is in place for all genomics institutes and all cloud computing providers. At GIS, we have had the most experience with Amazon Web Services (AWS). GIS has a collaborative project with American and European scientists to understand differences in cancer between Asians and Westerners. Collaboration is absolutely essential to tackle the complexity of cancer, yet the scale of the data led us to (briefly) consider shipping physical hard drives across the skies. Together with AWS, we solved the data transportation problem online, making it five times more cost-efficient and ten times faster, while simultaneously gaining flexibility to provide instant updates with new data. Indeed, digital communication trumps physical transportation.

First Steps Towards Genomics-Enabled Health

Can genomics in the cloud help us accelerate the solving of common problems in healthcare as well? Could we determine a person's risk for heart disease or cancer – within minutes? GIS, together with AWS, has taken first steps toward this possibility. While we can purchase perhaps a few thousand computers, now a single internet connection to AWS lets us command a fleet of millions. In the past, one of the critical steps in analysing a single person's genome used to take six hours; on AWS, this key step can be done in less than 15 minutes (and for a fraction of the cost).

Imagining the Future

Using movies as a snapshot of our imagination, we find it interesting that health is often a limiting resource. Money, transportation, communication are taken for granted; but our movie characters almost

always struggle with their health and their mortality. In the past decade alone, however, we've seen the surging transformation in genomics research; in another ten years, the impact of genomics will be as unrecognisable to us as mobile phones would have been to our grandparents when they were teenagers. Singapore is ideally positioned for this current phase of convergence between genomics and cloud computing, with the country's Smart Nation initiative driving forward transformation in all areas of technology, including healthcare. Genomics, combined with cloud computing, is truly bringing us to a future where our reality may even exceed our imagination.

GBS Callout

In Singapore in 2015, there was a large outbreak of infections associated with eating the local raw fish dish, *yu sheng*. The bacteria that caused the disease is called Group B Streptococcus, or GBS. No one had ever reported that GBS could cause infections through eating food; was this really happening in Singapore? GIS, part of the Agency for Science, Research, and Technology (A*STAR), partnered with other public agencies (Agri-Food and Veterinary Authority of Singapore, National Environmental Agency, Ministry of Health) and local hospitals (Tan Tock Seng Hospital, Singapore General Hospital, National University Hospital, Changi General Hospital) to investigate this outbreak using genomics. Together, we demonstrated that the same bacteria causing disease in humans was indeed found on the fish sold at the same hawker stalls they ate at! This was a remarkable finding, and this genomics result helped to prove that GBS does indeed cause severe infections due to eating contaminated food. This is a totally new mode of infection for GBS, which means we will need to change our medical teaching about GBS as well as change our education about food handling and food preparation.

Zika Callout

There was a Zika outbreak in Singapore in 2016. The major concern was whether this outbreak was imported from Brazil or elsewhere in the Americas, because that was where Zika was associated with birth defects (microcephaly) after mothers became infected. To answer this question, we used genomics, or sequencing of the viral RNA for Zika virus. In collaboration with other colleagues at Duke-NUS, Ministry of Health, and Environmental Health Institute, we adapted our expertise in Dengue genomics to study Zika samples from Singapore. We found that the Zika outbreak in Singapore was due to a strain of Zika that was from Southeast Asia – and not from Brazil or elsewhere in the Americas. This helped to provide information to the public health agencies for communicating the risks of Zika infection as well as to target mosquito control in Singapore (the same mosquitoes can transmit Dengue and Zika). The A*STAR investigators on this project were Paola Florez de Sessions (GIS) and Sebastian Maurer-Stroh (Bioinformatics Institute).