FINDING THE ACHILLES' HEEL OF PARASITES

Next-generation genomics in a wormy world – working toward radically new interventions

It was once a major challenge to sequence the entire genome of a pest organism to find clues for how to control it. Now, molecular parasitologists are using supercomputers to do just that – and they're working toward new ways of combating destructive parasites of humans and animals.

Written by Tim Thwaites, Science in Public

Barber's Pole Worm
Image credit: David Paul

vlsci
LIFE SCIENCES COMPUTATION

"Two years ago we were just dreaming about this"

Prof. Robin Gasser (top centre) and his team at Werribee, Victoria, Australia.
Image credit: David Paul

Using supercomputers and the latest gene sequencing technology, Professor Robin Gasser's research group from The University of Melbourne's Veterinary School is working hard to decode the genomes of a number of devastating parasites, in order to find new ways controlling them.

The researchers, who call themselves molecular parasitologists, employ the power of supercomputers to piece together the DNA and RNA of parasites into genomes, and compare them with other organisms. Once a parasite genome is assembled, it can be mined for useful and practical information.

The group began by working on the genome of a devastating sheep parasite, the Barber's Pole Worm (*Haemonchus contortus*), and have already developed, and are testing, a new treatment for it.

"We are completing a number of other parasite genomes at present," Robin says. "Two years ago we were just dreaming of this. It wouldn't have entered my mind to sequence a whole genome. Now we are doing it. The quantum leap in using supercomputers is really the ability to handle huge amounts of data quickly."

In their work, Robin's group used Bruce, the SGI Altix x86 cluster supercomputer at the Victorian Life Sciences Computation Initiative (VLSCI), to compare the Barber's Pole Worm's DNA and RNA with that of other organisms in order to track down genes essential for growth, development, reproduction and survival. Although these genes had not been detected in this worm before, they were able to be targeted by referring to what is already known about the genetics of related organisms such as the Free-Living Roundworm (*Caenorhabditis elegans*), the Fruit or Vinegar Fly (*Dropsohila*), yeast and mice.

Dr Neil Young, Research Scientist
Image credit: David Paul

In fact, the researchers were looking for specific enzymes in critical biological and metabolic pathways. Chemicals that act selectively to inhibit these enzymes were already known. Members of Robin's group not only found the enzymes for which they were searching, but they matched them to inhibitors. What's more, they were able to use their information to design new inhibitors with improved and specific properties which act upon worms.

But that's only part of the story. For the group has now moved on from projects involving single genes, to exploring whole genomes. "In general, we know very, very little about the genomics and genetics of parasites," Robin says.

Initially, researchers employed microarray technology to determine the genes which are active at different stages of the life cycle of a parasite. Then they sequenced this entire spectrum of active genes – in the case of the barber's pole worm, maybe 15,000 to 20,000 of them – using "next -generation" technology.

The sequencing of a whole genome, however, is much more complicated than simply assembling these active genes. A genome includes a lot of genetic material which does not seem to play any part in coding proteins at all. This is sometimes called "junk DNA". Some of it is involved with gene regulation; some of it is structural, and we have no idea of the role of much of it. Reconstructing the genome, including the junk DNA, is a bit like putting Humpty Dumpty back together again – a job for a supercomputer, indeed.

Much of the work so far has been undertaken together with expertise and facilities available at the Beijing Genomics Institute (BGI) in Shenzhen, near Hong Kong, and California Institute of Technology (Caltech). In addition, the group is undertaking a series of projects in collaboration with other researchers around the world. Then came high-throughput gene sequencing, which opened the way to studying whole genomes instead of working with single genes or short sequences.

But, with high-throughput gene sequencing another issue arose: data overload. "If you've got 300 million sequences and you pump them into a traditional computer, it's going to collapse/die. But, with supercomputers, we've got the grunt, the memory and the RAM available to start dealing with these datasets. They can handle all of the quality control, 'cleaning up', the filtering, the assembly, and making sense of the databases (annotation). It's a revolution."

But Robin's team of about five post-doctoral fellows and four doctoral students has had to extend to meet the challenge. "The work really requires close interaction between the biologists, who understand the organisms, and other scientists, including computational specialists."

And the researchers are breaking new ground as they go. The next step, for instance, is to prepare the data they are generating to be handled by the most powerful of the VLSCI computers, the IBM Blue Gene. "The other day we had a discussion with the guys from the IBM Collaboratory for Life Sciences based at VLSCI. They said that the problems we have with assembling the genomes of our parasites are quite similar to those they are facing with the assembly of some cancer genomes."

For further information about this research contact Professor Gasser at robing@unimelb.edu.au.

To contact VLSCI go to www.vlsci.org.au

vlsci
LIFE SCIENCES COMPUTATION

State Government Victoria
THE UNIVERSITY OF MELBOURNE
IBM

VLSCI is an initiative of the Victorian Government in partnership with The University of Melbourne and the IBM Research Collaboratory for Life Sciences, Melbourne. It exists for all Victorian researchers and aims to be one of the top five life science computation facilities in the world by 2013.