With help from Bioplatforms Australia, Australia has taken on chromosome 7A of the biggest genome in any cereal, looking for game-changing genes for antipodean growers.

By Graeme O’Neill

THE GREEN MONSTER
has kept its secrets hidden well over the millennia. Archaeologists know only approximately where and when bread wheat—the world’s No 3 food crop—was first cultivated. In separate genetic accidents somewhere within the 2300km arc of the Fertile Crescent, between 10,000 and 7000 years ago, the complete genomes of three diploid grasses collided and fused, giving rise to the progenitor of today’s highly productive bread wheats.

The complexity of bread wheat’s huge allohexaploid genome ensured it would be the last of the Big Three cereals—maize being No 1 and rice No 2—to have its DNA sequenced and its genes catalogued and mapped to its 42 chromosomes. The wheat genome project was a challenge of a different order. Dicing bread wheat’s A, B and D genomes and piecing them back together, when so many of their sequences are highly homologous, was always going to be a million bridges too far for shotgun sequencing. The wheat genome project demanded a new organisation, a new strategy, super-fast next-generation sequencing technology, and sufficient massively parallel computing capacity to stitch it all back together for mapping and annotation.

In 2005, a group of plant scientists, breeders and growers formed the International Wheat Genome Sequencing Consortium (IWGSC) to explore the structure and function of the bread wheat genome and lay the foundations for another quantum boost in wheat’s productivity, this time driven by molecular genetics. Professor Rudi Appels, a research leader in the Centre for Comparative Genomics at Murdoch University in Perth, is one of the consortium’s six international co-chairs.

Australia is the world’s second largest exporter of wheat, after the US, but Appels and Professor Peter Langridge, director of the Australian Centre for Plant Functional Genomics at the University of Adelaide, found it difficult to interest Australia’s major funding agencies in a basic research project with enormous potential, but no immediate payoff.

The wheat industry was whitener by years of severe drought, the Australian Wheat Board had been spruiking dubious dealings in Iraq, and the Australian Research Council declined to be involved. However, Appels and Langridge finally secured a grant from the Grains Research and Development Corporation.

The IWGSC was assigning individual chromosomes from wheat’s A, B and D genomes to member nations, relying on the expertise of Professor Jaroslav Dolezel’s cytogenticists and cytometry team at the Institute of Experimental Botany in Olomouc, in the Czech Republic, which has mastered the nanoscale art of sorting and dissecting out individual chromosomes from wheat cells.

“We put our hands up for chromosome 7A. We were able to convince the GRDC to fund it, not because of its importance as a fundamental entity, but because it contains some important genes that influence yield quality and disease resistance,” says Appels. “And Bioplatforms Australia agreed to come on board, so Australia was now making a pretty significant contribution to the project.”

BIOPREPARATIONS ON BOARD
Bioplatforms Australia was established under the Commonwealth Government National Collaborative Research Infrastructure Strategy as a one-stop-shop of genomics, proteomics, metabolomics and bioinformatics databases for Australian bioscience researchers.
"As soon as you get into the genome, it's like being in outer space. There are no restrictions on where we can go and what we can do."

Bioplatforms Australia assigned Dr Anna Fitzgerald to run the project, and liaise with wheat researchers.

The company is running separate but complementary projects on wheat and wheat's major pathogens, and relying on researchers like Dr John Manners, deputy chief of CSIRO Plant Industry, and an expert in the genetics of host-pathogen interactions in wheat, to link the two fields.

Fitzgerald expects it will take three to four years to complete chromosome 7A. "The technology is changing and we can now consider carrying out the full sequencing of the 8000 BACs (Bacterial Artificial Chromosomes), used to store and clone large DNA sequences) that define the entire chromosome."

The donor wheat for chromosome 7A is Chinese Spring Wheat, which has been designated as the reference genome for the international project. Bioplatforms Australia, working with Australian breeders and wheat researchers, has also witnessed a list of nearly 60 Australian varieties down to 23, based on their genetic importance in different regions of Australia, and notable attributes or productivity in different environments.

Elite varieties like Western Australia's prolific Wyalkatchem and eastern Australia's Vixi will undergo genomewide association studies based on new genome sequence information to identify markers associated with yield, drought tolerance and disease resistance.

The time will be done with the Multi-parent Advanced Generation Inter-Cross (MAGIC) breeding lines developed by CSIRO's Food Futures Flagship. The MAGIC lines were bred to bring together gold-standard genetic variation for bread-baking quality, water absorption, milling yield and protein content from Australia's best varieties, in as few lines as possible, facilitating genetic analyses for these traits.

Fitzgerald says the substantial homology across the A, B and D genomes makes hexaploid wheat a challenge. Much of the study will rely on the reference genome for chromosome 7A and the other wheat chromosomes as a templates, and as other countries around the world start to produce data to define the DNA sequence detail of the wheat genome.

"All of these programs will involve direct collaboration between Bioplatforms Australia and the research community," says Fitzgerald. "We need to know what sort of datasets would help address the critical challenges facing researchers - they will prioritise the targets, we will generate the sequence and support the bioinformatics effort required to pull it all together. We'll need joint investments to bring the different datasets together at the end, and make them usable to a range of different researchers with varying interests."

**PATHOGEN Datasets**

Fitzgerald said wheat pathogen researchers want datasets for multiple strains of the Australian wheat industry's high-risk pathogens. *Puccinia stem, flag and leaf rusts*, including strains of Africa's virulent Ug99 stem rust, the most direct threat to Australia's wheat industry, feature high on the most-wanted lists, along with *Stagonospora blotch* and *Pyrenophora zunt*. Then says. "A lot of work has been done already, and our work will complement it, and help fill the gaps."

"The GRDC has provided strong backing for this work, and we're working with them to ensure the research that can be done off the back of these datasets can be funded. There's no point producing data if it's just going to sit there. We're assembling the datasets so they can feed directly into the breeders' work."

Fitzgerald says Bioplatforms Australia is investing around $2 million of its own funds, but trying to quantify the total cost would be for the project. This goal will need to be raised, but Fitzgerald said the company is confident it can meet the challenges of sequencing the wheat genome.
of the project is difficult due to the breadth of the research organisations involved in the project and the multiple sources of funding each organisation depends upon.

Appels expects Australia's participation will attract top people from the next generation to do cutting-edge work. "With rice and maize completed, it's not easy to attract the young minds to research wheat when we don't even have a genome project organised.

"For growers, the targets are yield, yield and yield, but a lot of novel aspects of yield are only now emerging from studies of model organisms like rice. It turns out that in rice, the grain actually represses itself so it doesn't grow too large. If we can release that repression in cultivation, we can substantially increase grain size. We'd like to capture that effect in wheat, and we have already identified some genes that limit search production by controlling cell division in endosperm."

According to Appels, there are many other potential targets in wheat that promise to yield new germplasm for breeders to deal with traits like frost resistance, drought tolerance and other environmental stressors, all of which contribute to yield losses.

For more than a century, classical wheat breeders have battled the sheer randomness of the meiotic lottery, and linkage disequilibrium - the tendency of particular combinations of alleles of genes on the same chromosome segment to time-travel together as extended haplotypes - so that new breeding lines inherit much dross with the gold.

"We now have genetic modification technology moving in wheat, although it's still largely inefficient," he says. "So we can target specific genes, divide and conquer, and achieve yield gains without dragging along unwanted genes that detract from crop quality, or vice versa.

"As soon as you get into the genome, it's like being in outer space. There are no restrictions on where we can go and what we can do. For example, we can go into familiar cultivars with high flour quality, and release their yield repressors, which is what interests the GRDC. They're willing to back us in tracking down the more special areas, because it's going to give us some really big breakthroughs."

SNPS AND CHIPS

Appels is delighted that the GRDC is investing in a big project with his former CSIRO colleague, Dr Matthew Morell, now with the WA Department of Agriculture, which is targeting grain quality. The GRDC was more amenable to investing in targeted genome projects in traits like yield and quality because it improved the odds of big payoffs for commercially important traits.

"The real benefit for breeders: that we'll come out of the whole genome effort is the ability to detect single-base DNA differences between varieties. In that respect, wheat is an absolute monster to deal with because it's very hard to find the differences that, for example, make Wyal Creek so different from other cultivars. Those differences might be driven by single-nucleotide polymorphisms (SNPs) in only one or two genes, but you have to do a whole genome-wide association study to find them."

"So Matthew Morell, Matt Hayden in Victoria and a group in CSIRO are working with a group in Kansas to develop a SNP chip, the cutting edge technology for detecting and mapping SNP variation across many lines. They currently have a 9000-SNP chip, and next year they hope to have a 50,000-SNP chip.

"With a 50,000-SNP chip, and a 50 per cent hit rate, we would expect to find at least 20,000 SNPs that differ between Wyal Creek and other varieties. Many researchers already have that capability, but it's still a bit expensive for wheat."

Appels expects an early draft of TASA should be completed by the end of this year, but much work will then be required to link it all together. Bioplatforms Australia's brief is to incorporate all the international sequence data in its databases.

"Where the real power comes, and the GRDC really starts to pay attention, is when we can anchor all that molecular biology to a genetic map, because then we can link the SNPs to the quantitative trait loci (QTLs) you can touch and feel in the field, that have already been located on the map, and look for the genes responsible for the traits we're interested in.

"This is a really challenging problem in bioinformatics. Matthew Belgradation at the Centre for Comparative Genomics Murdoch University, is running that arm of the program for Bioplatforms Australia, to set up the databases that carry a huge volume of easily searchable data to make the linkages. It's a job for specialists, not wheat breeders or gene jockeys.

"As an example, if I have variety X, and I know about its yield and phenotype, and the chromosome regions with an unidentified gene of interest, I can then go the Bioplatforms Australia databases and run a candidate-gene analysis, based on what we already know from rice. I narrow it down to a suite of half a dozen genes, validate them and go to the breeders and tell them: 'I believe these polymorphisms are all key contributors to yield in your environment'."

"The metabolomics comes in at the other end of the scale, providing super high-definition resolution of the phenotype, in terms of molecules we can isolate and define. We can see a change in the metabolite that correlates with a change in phenotype, which takes us back to the genes.

"Without fundamental, well-organised databases, we'd really be stuck."

NEW SureSelect Kits: Mouse Exome | RNA Capture
Validated Automation Protocols | Additional Plus kits (All Exon v2, 50Mb and Kinome)

Agilent's SureSelect is the world's leading target enrichment system. It's the most proven, flexible, and comprehensive up-front capture solutions for next-generation sequencing.

• BETTER FLEXIBILITY

100% compatible with the three leading NGS platforms.

• BETTER PUBLISHED

More peer-reviewed publications than all other target enrichment solutions combined.

• BETTER SCALABILITY

Meets the throughput needs of any size study.

• BETTER PERFORMANCE

Delivers unsurpassed allele detection reliability; reliably captures SNPs, indels, and CNVs.

The Measure of Confidence
Agilent Technologies

Do the science you want to do—the way you want to do it.
Get SureSelect, www.agilent.com/genomics/sureselect

© Agilent Technologies, Inc. 2011