Overcoming agricultural and human health challenges with KASP genotyping chemistry

BIOSEARCH™ TECHNOLOGIES

GENOMIC ANALYSIS BY LGC

Contents

Introduction

As a growing population challenges the world's limited resources, scientists around the globe are leveraging sophisticated genomics technologies to improve the yield, quality and sustainability of our food supply. But despite the wealth of genome sequence information that already exists, a tremendous amount of work remains to be done to convert the sequence into a usable format, work that ties specific traits, phenotypes and functions to genes and genome regions.

At LGC Biosearch Technologies™, we develop and commercialise enabling genomics technologies, such as our proprietary KASP™ genotyping chemistry, that make powerful genomics approaches efficient and affordable enough to conduct at industrial scales. With KASP, teams in Colorado, USA, and the United Kingdom are breeding hardier strains of wheat, an oil palm plantation in Malaysia is increasing the sustainability of their operations and plant biologists in California, USA, are developing disease-resistant peppers.

In this collection, we highlight just a few of the ways KASP genotyping chemistry is making a difference in the world, helping scientists achieve mission critical projects aimed at overcoming today's many food sustainability and health challenges. We hope these stories inspire you to consider how KASP technology can support your efforts to address these global challenges.

Working together with you, Biosearch Technologies is using science for a safer world.

Let's get started with how KASP technology works followed by a deep dive into each individual success story.

In this collection, we highlight just a few of the ways KASP genotyping chemistry is making a difference in the world

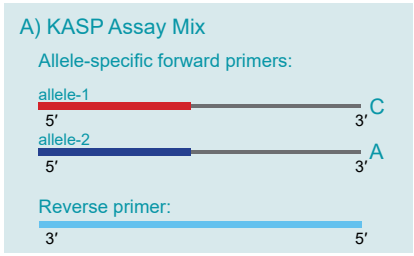
How KASP works

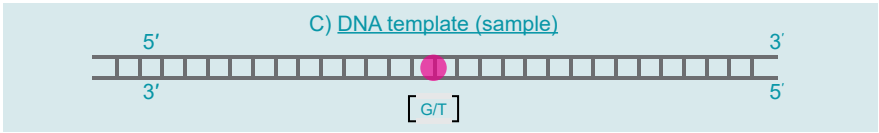
1. Assay components:

A) KASP Assay Mix: consists of 2 allele specific primers and 1 reverse primer.

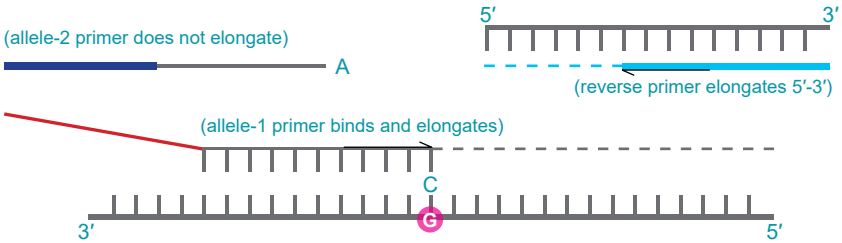
B) KASP Master Mix: contains universal fluorescent probes, Taq polymerase and dNTP's in an optimised buffer solution.

C) Sample DNA: DNA contains the SNP of interest.





2. Denatured template and annealing components – PCR round 1:

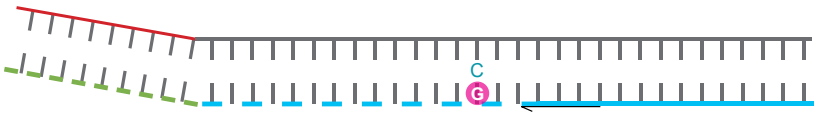


In the first round of PCR, one of the allele-specific primers matches the target SNP and with the common reverse primer, amplifies the target region.

Legend

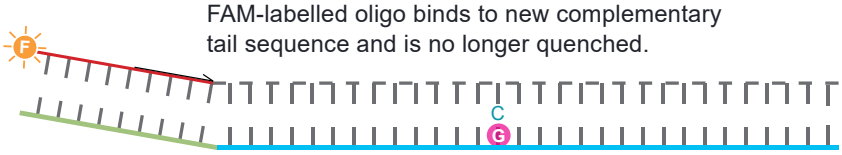
- Allele-1 tail FAM-labelled oligo sequence
- Allele-2 tail HEX-labelled oligo sequence
- Common reverse primer
- F — FAM dye
- H — HEX dye
- Target SNP
- Q — Quencher

3. Complement of allele-specific tail sequence generated – PCR round 2:

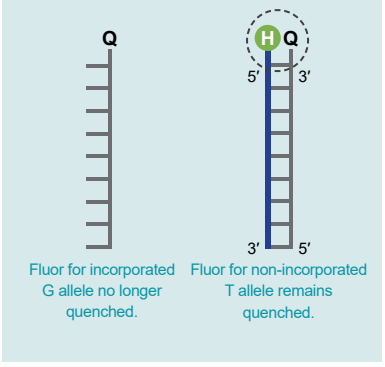


(Reverse primer binds, elongates and makes a complementary copy of the allele-1 tail.)

4. Signal generation – PCR round 3:



In further rounds of PCR, levels of allele- specific tail increase. The fluor labelled part of the FRET cassette is complementary to new tail sequences and binds, releasing the fluor from the quencher to generate a fluorescent signal.



Application note

Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat *(Triticum aestivum L.)*

Keith J. Edwards, University of Bristol, Bristol, UK; LGC Biosearch Technologies™, Hoddesdon, UK

Summary

Wheat is a key food source both for humans and animals, but the modern varieties are reducing in genetic diversification. Introducing genes from related species could increase its resistance to pests and diseases, and increase crop yields. Improvements in next generation sequencing (NGS) and our understanding of wheat's complex genome will help the process to identify molecular markers for useful wheat characteristics, to improve this development of novel wheat cultivars.

Kompetitive Allele Specific PCR (KASP™) genotyping technology is a one-step system that can validate SNPs (single nucleotide polymorphisms) that have been identified by NGS, and is particularly suited to use with polyploid species. Researchers led by the University of Bristol used KASP to validate varietal SNPs mined from UK and European cultivars of wheat. The SNPs were derived from a publicly-available wheat expressed sequence tag (EST) database and unique sequences generated by the team using NGS.

The team created the first large scale assembly of genotyping and genetic map information for elite UK wheat varieties based on individual SNP markers, and was the first to demonstrate the use of KASP-based technology to genotype wheat cultivars and generate an extensive linkage map. The research goes on to find more SNPs in areas where the information is sparse.

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC

Application note

Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat *(Triticum aestivum L.)*

Introduction

The global population is growing, and it is estimated that it will reach nine billion by 2050. There is an increasing demand for cereals, both as a food source and as feed for animals[1]. As wheat is among the three most important food sources and the main cereal crop for temperate climates[2], increasing the yield of wheat crops will be very important[3].

Modern bread wheat has a complex genome; it is allohexaploid (AABBDD), formed by introgression between the diploid (DD) genome of *Aegilops tauschii* and the tetraploid (AABB) genome of *Triticum turgidum*[4]. Modern wheat has lost a large proportion of its genetic diversity[5], and introducing genes from its wild relatives could increase stress, pest and disease resistance[6]. Finding and using molecular markers such as SNPs will enable precision breeding of important desirable traits.

This application note describes the use of KASP genotyping technology from Biosearch Technologies to validate varietal SNPs mined from UK wheat germplasm[7].

The map positions of the different SNP marker types were investigated to determine whether any bias in genetic location was introduced by using co-dominant SNP assays in two double haploid (DH) mapping populations developed from UK cultivars Avalon x Cadenza (A x C) and Savannah x Rialto (S x R). Of the 3214 SNP markers developed to date, 2109 were identified as polymorphic between Avalon and Cadenza, of which 1807 were placed on the Avalon x Cadenza map.

‘KASP genotyping technology is a one-step system that can validate SNPs that have been identified by NGS, and is particularly suited to use with polyploid species.’

Application note

Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat *(Triticum aestivum L.)*

Materials and methods

Plant material

For SNP genotyping, the researchers selected 23 different wheat varieties to represent the UK wheat germplasm, harvesting and freezing the root and leaf tissues at six weeks.

Preparation of normalised cDNA libraries

For NGS, the researchers extracted and purified total RNA from root and leaf tissue of five wheat varieties (Avalon, Cadenza, Rialto, Savannah, Recital) and then synthesised, purified and normalised complementary DNA (cDNA).

Next generation sequencing

The University of Bristol Transcriptomics Facility processed and sequenced 5 µg of normalised cDNA for each variety.

SNP discovery

Putative varietal SNPs came from two sources:

- The publicly-available wheat expressed sequence tags (EST) data from NCBI (National Center for Biotechnology Information) – the researchers selected 213 SNPs with a high probability of being varietal for validation.

- Combining the NCBI wheat sequence data set with the next generation sequencing data from the five wheat varieties sequenced as part of the study. The resulting reference data set of 91,368 sequences was mined for SNPs with at least two alternative bases at a reference position.

SNP validation

For each putative varietal SNP, KASP assays (two allele-specific forward primers and one common reverse primer) were designed by Biosearch Technologies. Genotyping reactions were performed in a thermal cycler in a final volume of 5 µL, containing 1x KASP Master Mix, the SNP-specific KASP Assay Mix and 10-20 ng of genomic DNA.

Importantly, the cycling conditions were the same for every SNP assay:

- 15 min at 94 °C
- 10 touchdown cycles of 20 sec at 94 °C
- 60 sec at 65-57 °C (dropping 0.8 °C per cycle)
- 26-35 cycles of 20 sec at 94 °C
- 60 sec at 57 °C

After detecting the fluorescence and analysing the data, the team calculated polymorphic information content for each marker, and carried out hierarchical cluster analysis.

Genetic map construction

The software program MapDistro v. 1.7 was used to place the SNP markers into the previously established genetic map derived from 190 Avalon x Cadenza doubled haploid lines. The Kosambi mapping function was used to calculate map distances.

Results and discussion

This was the first large scale assembly of genotyping and genetic map information for elite UK wheat varieties based on individual SNP markers.

The mining for SNPs specifically targeted varietal SNPs in exon sequences in order to generate molecular markers with a potential link to QTL's, and to include both synonymous and non-synonymous point mutations.

Application note

Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat *(Triticum aestivum L.)*

SNP discovery

By screening the publicly-available wheat EST database the researchers identified around 3500 putative varietal SNPs in 8668 sequences. Sequencing the normalised whole-seedling cDNA from the five wheat varieties grown for the purpose generated 24-45 million, 75-base paired-end reads for each line, and SNP discovery led to 14,078 putative SNPs in 6255 distinct reference sequences (2.7 megabases). This is an average of five varietal SNPs per kilobase in the reference sequences that have one or more SNP.

Both techniques identified a similar number of varietal SNPs per kilobase, and matched previous work on varietal SNPs in wheat[9; 10].

SNP validation and characterisation

The researchers selected 1659 putative SNPs for validation with genomic DNA (213 from sequences from the NCBI database; 1446 from the NGS process). The NGS SNPs were selected based on their predicted polymorphism level and validated using the KASP genotyping platform on 21 hexaploid wheat varieties, with a diploid and a tetraploid variety.

Of these 1659 SNPs, 1114 (67%) were polymorphic between the different varieties. This conversion rate of around 67% from the KASP system is relatively high for a complex polyploid species[11], but still has potential to be increased using primer design and PCR optimisation if necessary. 70 (4%) were monomorphic in the hexaploid varieties, but polymorphic between the hexaploid varieties and the diploid and/or tetraploid, which suggests that these markers represent intravarietal homoeologous SNPs.

The primers can be redesigned as homoeologous-specific KASP primers, to discriminate when homoeologous copies do or do not contain varietal SNPs. Screening markers from this category against the Chinese Spring nullisomic lines confirmed that this was the case in 26 of 28. As intravarietal homoeologous SNPs normally account for around 74% of all SNPs in wheat[10], this suggests that these data sets have been enriched for varietal SNPs. However, as the process is time- and work-intensive, this is probably only useful for specific projects, such as investigating homoeologs tightly linked to specific loci of interest.

475 (29%) failed to generate a useful amplification signal, and this could be improved by using optimised PCR conditions and alternative primers.

Application note

Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat *(Triticum aestivum L.)*

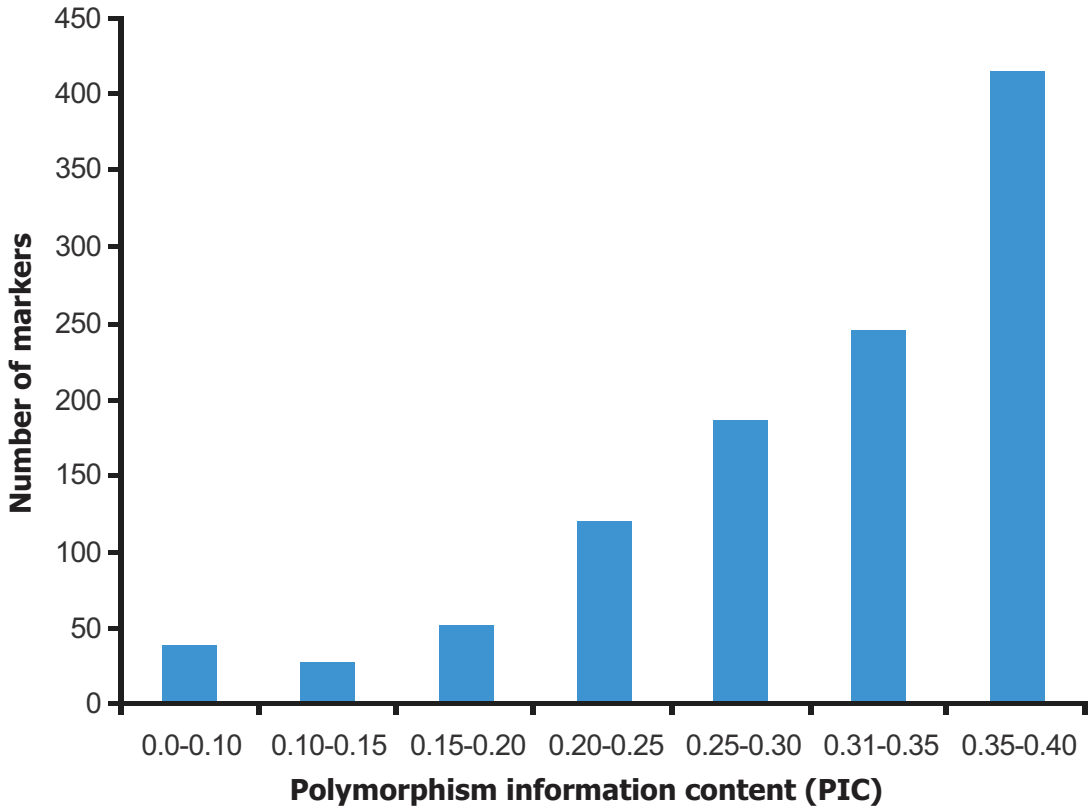


Figure 1. Distribution of PIC of SNPs among the 23 wheat varieties.

The polymorphism information content (PIC) values for the validated markers were 0.08-0.975, with an average of 0.300, comparable to other results from wheat[11; 12], and no significant differences between markers from different genomes or homoeologous groups (Figure 1).

Application note

Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat *(Triticum aestivum L.)*

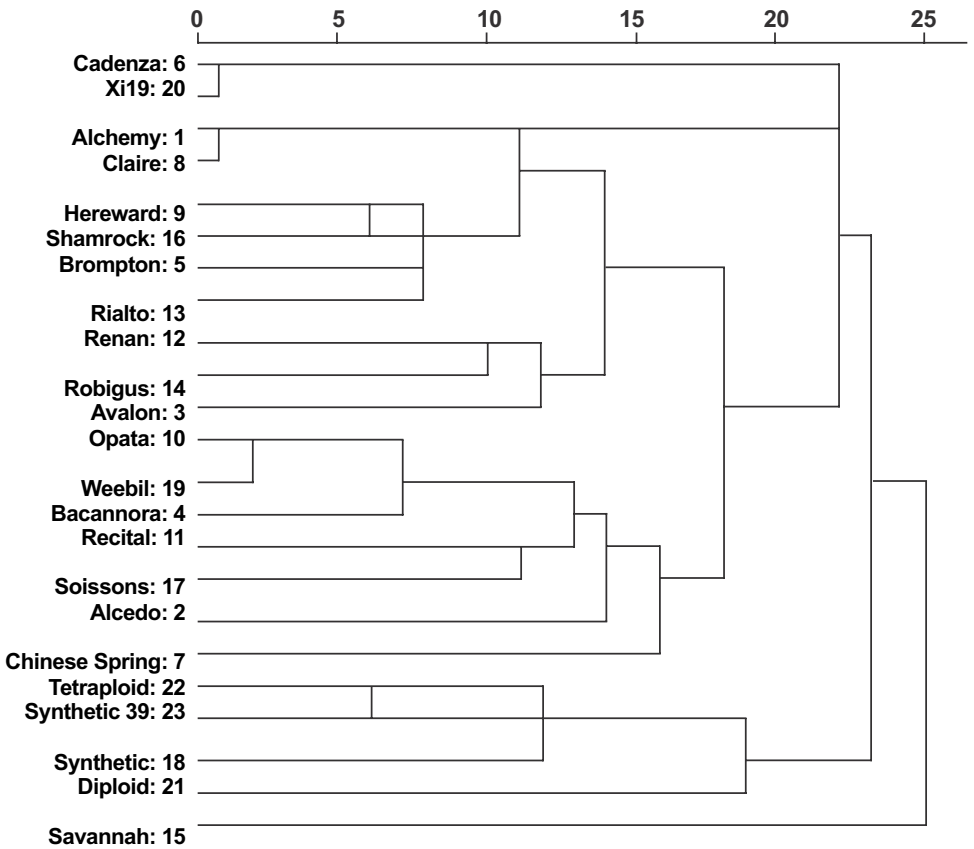


Figure 2. Genetic relationships between the 23 wheat lines screened with 1114 KASP probes.

Hierarchical cluster analysis suggests possible genetic relationships between the 23 lines (Figure 2).

Application note

Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat *(Triticum aestivum L.)*

Genetic map construction

This study produced the first SNP-based linkage map, based on 500 transcript-linked varietal SNPs and 574 existing markers. The total map length is similar to maps previously created[13; 14; 15] using a variety of markers, and showed similar map lengths for the A, B and D genome. There was a significant difference in the distribution of markers between the three genomes, and the markers were not evenly distributed between the homoeologous chromosome groups, with markers clustered around the centromeres, particularly on the A and B chromosomes.There were fewer SNP markers on the D chromosomes, both compared with the A and B chromosomes and with other marker studies. This is probably because SNP markers target genic regions, and this reflects the lack of genetic diversity in modern wheat varieties. Future SNP discovery studies should target the D chromosomes to counterbalance this lack of diversity[12; 16].

The researchers compared the genetic map locations of the 500 SNP loci with their predicted map location based on physical mapping experiments. Overall, the correlation was relatively high, and is useful for identifying and mapping deletions. Discrepancies, for example SNPs mapping further apart than expected based on linkage group information, suggests that the relative positions of ancestral markers have been altered through chromosomal rearrangements.

Some SNPs were similar to sequences in the NCBI non-redundant protein sequence database but mapped to multiple regions of the genome. This suggests that KASP-based genotyping has high enough levels of sensitivity to map individual paralogous genes in a complex polyploid genome such as the wheat genome.

Conclusion

According to the researchers, this is the first report of a public linkage map for hexaploid wheat containing several hundred individual SNP markers, and the first demonstration of KASP-based technology to genotype wheat varieties and generate a linkage map.

Creation of the linkage map took 102,220 individual KASP reactions, using 538 probes on 190 plants.The reactions were carried out within 24 hours, using simple microplate technology, and read with a standard fluorescence resonance energy transfer (FRET)-capable plate reader. The team behind the study is confident that this could lead to fast and cost-effective genotyping of thousands of plants with a large and flexible number of markers.

The next step will be to find more SNPs, particularly focusing on the D genome and the homoeologous group 4 chromosomes, which would make genome-wide association studies (GWAS) possible in wheat. This project has also shown that it‘s not just the SNP markers that are important, but also the information associated with these, such as the surrounding sequences. This additional information means that the SNP data can be used in a variety of genotyping platforms, now and in the future. Making this information freely available will provide the power needed to develop new varieties of wheat and meet the global food needs in the future.

Application note

Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat *(Triticum aestivum L.)*

References

1. Foresight. The Future of Food and Farming: Final Project Report. s.l. : The Government Office for Science, 2011.
2. Wheat. Shewry, P R. 2009, J Exp Bot, Vol. 60, pp. 1537-1553.
3. Reaping the benefits: Science and the sustainable intensification of global agriculture. s.l. : The Royal Society, 2009.
4. Genome plasticity a key factor in the success of polyploid wheat under domestication. Dubcovsky, J and Dvorak, J. Jun 2007, Science, Vol. 316, pp. 1862-1866.
5. Grinding up wheat: a massive loss of nucleotide diversity since domestication. Haudry, A, et al. Jul 2007, Mol Biol Evol, Vol. 24, pp. 1506-1517.
6. Raising yield potential of wheat. I. Overview of a consortium approach and breeding strategies. Reynolds, M, et al. Jan 2011, J Exp Bot, Vol. 62, pp. 439-452.
7. Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (Triticum aestivum L.). Allen, A M, et al. Dec 2011, Plant Biotechnol J, Vol. 9, pp. 1086- 1099.
8. Sambrook, J, Fritsch, E F and Maniatis, T. Molecular Cloning: A Laboratory Manual. 2nd edition. Cold Spring Harbor : Cold Spring Harbor Laboratory Press, 1989.
9. DNA sequence polymorphisms and their application to bread wheat quality. Ravel, Catherine, et al. 3, December 2007, Euphytica, Vol. 158, pp. 331-336.
10. A genome-wide analysis of single nucleotide polymorphism diversity in the worlds major cereal crops. Barker, G L and Edwards, K J. May 2009, Plant Biotechnol J, Vol. 7, pp. 318-325.
11. Multiplex single nucleotide polymorphism (SNP)-based genotyping in allohexaploid wheat using padlock probes. Edwards, K J, et al. May 2009, Plant Biotechnol J, Vol. 7, pp. 375-390.
12. Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (Triticum aestivum L.). Chao, S, et al. 2010, BMC Genomics, Vol. 11, pp. 727-727.
13. A high-density microsatellite consensus map for bread wheat (Triticum aestivum L.). Somers, D J, Isaac, P and Edwards, K. Oct 2004, Theor Appl Genet, Vol. 109, pp. 1105-1114.
14. Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. Akbari, M, et al. Nov 2006, Theor Appl Genet, Vol. 113, pp. 1409-1420.
15. A high-density genetic map of hexaploid wheat (Triticum aestivum L.) from the cross Chinese Spring x SQ1 and its use to compare QTLs for grain yield across a range of environments. Quarrie, S A, et al. Mar 2005, Theor Appl Genet, Vol. 110, pp. 865-880.
16. Sequence polymorphism in polyploid wheat and their d-genome diploid ancestor. Caldwell, K S, et al. Jun 2004, Genetics, Vol. 167, pp. 941-947.

BIOSEARCH™ TECHNOLOGIES

GENOMIC ANALYSIS BY LGC

By tracking SNPs they are able to link specific genes to desirable crop characteristics or 'traits'. When a gene of interest is found, the plant carrying this gene is then backcrossed into the population in the greenhouse to fully isolate it or 'clean it up'.

Customer case study

Colorado State University Wheat Breeding and Genetics Program

The Colorado State University Wheat Breeding and Genetics Program is a long-standing initiative run out of Colorado State University (CSU), USA. Lead by Scott Haley, Professor of Plant Breeding at CSU, this ground-breaking program is using traditional breeding techniques combined with modern plant genomics to breed the next generation of wheat varieties for the Colorado wheat growing industry and beyond. Their locally-developed winter wheat varieties are specifically adapted to Colorado's tough climate.

Over 60% of all the wheat grown in Colorado is supplied by this wheat-breeding group, as well as Kansas, Nebraska and Montana, and there is a very close symbiosis between them and the growers they supply. Scott Haley and his team also conduct basic research into genetic and environmental factors that affect wheat yield and end-use quality.

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC

The work undertaken by Scott Haley and his team of extremely dedicated and experienced research associates involves:

- Plant genomics research
- Greenhouse propagation
- Double haploid production
- Screening
- Extensive field testing.

This combined effort produces new wheat varieties specifically bred, selected and tested for yield, drought tolerance, disease resistance, quality and other essential traits.

Using molecular biology for plant breeding:

Scott Haley has used molecular biology as part of his wheat-breeding program at CSU for many years, in particular marker-assisted selection (for intergeneric introgressions) and wheat-maize doubled haploid development (for breeding line development and parent building).

Scott Haley's research team sample the leaf tissue, extract the DNA and then genotype the samples in order to identify and track single nucleotide polymorphisms (SNPs) within a set wheat population. By tracking SNPs they are able to link specific genes to desirable crop characteristics or 'traits'. When a gene of interest is found, the plant carrying this gene is then backcrossed into the population in the greenhouse to fully isolate it or 'clean it up'.

"From October to early June we are gene jockeys, working in the lab, generating and analysing data; from mid-June to early October we put our overalls on and become farmers. It is a very holistic process".- Tori Anderson, Research Associate on Scott Haley's team.

The team takes advantage of an extensive network of local trial locations. These trial locations include eighteen main field plot locations around Colorado and west Kansas, over 17,000 wheat plots and 30,000 early generation head rows where the team carry out pilot studies and initial analysis.

The challenge: time and money

Narrow windows between harvest in July and planting in September in the University's wheat breeding programme schedule mean that the programme's researchers need to process a large number of samples for genotyping with a very quick turnaround.

Initially the team were using assays for SSR markers but discovered that SNPs were much more plentiful.

The team has long collaborated with the USDA Hard Winter Wheat Genotyping laboratory in Manhattan KS for routine marker analyses. These tests have included SSR markers and more recently KASP assays for SNP marker's linked to key traits in the breeding programme. The staff at the USDA has been especially helpful in optimisation of new markers for selection and sharing information on

development of new KASP markers.

As part of her work on Scott's team, Tori Anderson investigates new genotyping technologies to integrate into their wheat breeding programme. The group started looking into KASP genotyping a couple of years ago, when it was mentioned by the USDA genomics facility, who were talking about using KASP markers for their work.

A TILLING program at CSU is working to identify novel variants for traits of interest in wheat production or utilisation. Sanger sequencing allowed a greater degree of customisation but they found that sequencing in order to genotype every SNP mutant that they wanted to look into was very expensive and took time.

The group needs to perform SNP assays on a large number of samples quickly to allow them to select variants. In the back crossing programme this is normally a one month window, but can be more or less depending on the wheat strain they are growing, how fast it grows and greenhouse conditions.

Discovering KASP

Tori Anderson contacted our genomics technical support team to find out what we could do for them and found that KASP assays lend themselves very well to exactly the type of work that CSU is doing. KASP genotyping allowed assays for new SNPs identified by the team at CSU to be designed and made available very quickly; they have had great success with how easy KASP genotyping is to use and are now trying to convert as many markers as possible to KASP assays.

KASP SNP genotyping is based on a universal detection system where the fluorescence detection system is in the KASP Master Mix and generation of an assay against a new target is only a matter of ordering and assembling three short unlabelled oligonucleotides designed using the software of LGC Biosearch Technologies™. For each new SNP marker, new KASP assays can easily be designed, ordered and supplied at very little cost.

After an initial period trialling KASP genotyping to see whether it would work for their projects, the team has now been using KASP in earnest for full blown production work for over a year. Adoption of KASP genotyping has enabled the team to do more where previously costs and time prohibited them.

They order new KASP assays as KODs (sometimes KBDs) and run them in their lab on BIO-RAD CFX real-time machines. In this way the team in the past year conducted over 16,000 total KASP assays to enable selection for 15 different trait-associated SNPs.

The team started with simple SNP assays, representing novel mutations that they want to track, and are now working with more complex assays analysing small and now larger indels (insertation/deletions) as well as haplotypes.

Desirable traits – human health and pest resistance

As well as breeding crops tailored for the Colorado climate, traits for human health are now a focus, including traits related to starch production and antioxidant properties. Another example is selection for strains of wheat with specific aleurone layer colour, which could potentially improve product color in whole grain products made from hard white wheat. Consumers like to have the health benefits of wholewheat flour without the brown colouring that it gives to the bread, and improved strains allow the more of the healthy bran to survive the milling process without carrying with it the brown colour.

Wheat produces a volatile compound, which is attractive to crop-damaging pests such as the wheat stem sawfly. Wheat varieties are being selected, that produce less volatiles and are therefore less likely to be attacked by pests.

Funding

The Wheat Breeding and Genetics Program is funded in a very unique way. The programme has a very strong relationship with wheat industry groups in Colorado and the majority of funding for their work comes from the Colorado Wheat Administrative Committee (CWAC) and the Colorado Wheat Research Foundation (CWRF). These funds are used to complement and leverage the vital base support that comes from the university through the Agricultural Experiment Station (the US Land Grant University funding mechanism for agricultural research).

A small percentage from every bushel of wheat, grown by the varieties CSU supplies, is paid back into the group to fund their work. All or parts of the salaries of 5 of the programme's workers are paid in this way, plus laboratory instruments and consumables.

"They fund our work and our equipment and in return we are able to work better…to produce a better product for them. It is a real symbiotic relationship". - Tori Anderson

This unique funding mechanism gives the group freedom to explore and find the technologies that work, but with that responsibility to deliver for their sponsors. The consortium also funds postdoctoral positions which have allowed the team to look more in depth at double haploid wheat varieties, TILLING mutants and genomic selection.

'In our sixth decade of continuous wheat breeding activity, we continue to enjoy the excellent support provided through the partnership between CSU and the wheat industry in Colorado. Funding for our program comes from a combination of state and federal funds provided by the Colorado Agricultural Experiment Station (CSU AES) and Colorado wheat industry groups, including the Colorado Wheat Administrative Committee (CWAC) and the Colorado Wheat Research Foundation (CWRF).'[1]

CSU developed wheat varieties

- Since inception of the CSU's Wheat Breeding and Genetics Program in 1963, over 35 improved wheat varieties have been developed and released by the program
- Over 60% of the Colorado wheat acreage is planted to CSU-developed wheat varieties.
- Colorado growers have made substantial yield gains with new CSU varieties like Hatcher, Ripper and Bill Brown, proven in 2010, when Colorado set a new yield record of 45 bushels per acre. Recent releases such as Snowmass, Byrd, Brawl CL Plus, and Denali promise to bring even greater benefits to the wheat industry in Colorado and surrounding states.
- Varieties are marketed under the PlainsGold Brand.
- Over $3 million has been returned to CSU under the CWRF royalty program. This is money that will then fund further improved varieties for the Colorado growers.

Reference

1. Wheat Breeding and Genetics Program; https://wheat.agsci.colostate.edu/; Accessed November 2013.

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC

It has been estimated that the demand on cereal production will increase by 50% by 2030; scientists and breeders alike must now look to understanding the genetic make-up of wheat to help maximise its growing potential.

Customer case study

University of Bristol – wheat genomics

Why wheat?

Wheat is one of the three most important crops for human and livestock feed, and with food supply an increasing global concern, the pressure is on to increase cereal crop production as a solution to feeding the growing world population.

It has been estimated that the demand on cereal production will increase by 50% by 2030; scientists and breeders alike must now look to understanding the genetic make-up of wheat to help maximise its growing potential.

Industry challenge: struggling to meet a growing demand

Compared with other crops, the increase in wheat yields has slowed since the 'green revolution' of the 20th century due to the domestication of the crop.

This domestication of wheat has resulted in a decline of genetic diversity and it has been suggested that in order to increase it, genes from 'wild relatives' of the crop could be introduced. Such strategies, often referred to as pre-breeding, can be resource intensive.

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC

Diploid crops (consisting of two sets of chromosomes), such as rice and barley, have benefited from extensive genetic analysis and molecular breeding programmes. The complexity of a polyploid species (more than two sets of chromosomes) such as wheat coupled with a decline in genetic diversity has hindered these types of studies.

Aside from the challenge of meeting a growing global demand for wheat, the crop is also vulnerable to pests and disease. For example, the stem rust pathogen known as Ug99 is one of the most serious diseases found in wheat and poses a significant challenge to production. To combat this, researchers need to identify genes within the wheat genome that can provide a strong and durable resistance; access to new screening technology is essential in defending it from this and other threats and to sustainably produce sufficient and safe food.

Technical challenge

The Cereal Genomics group, based in The School of Biological Sciences, University of Bristol, has recognised the need to better understand the complexity of the wheat genome in order to assist molecular breeding programmes.

Marker assisted selection (MAS) is a process whereby a marker (morphological, biochemical or one based on DNA/RNA variation) can be tracked through populations and used for indirect selection of genetic determinants of the trait of interest such as:

- Productivity
- Disease resistance
- Stress tolerance
- Quality.

MAS has a long history of successful use by academic laboratories, genotyping service providers and breeding companies for this type of genetic analysis in plant and animal breeding.

However, the complexity of the wheat genome coupled with a decline in genetic diversity within modern elite cultivars has made application of MAS in wheat breeding programmes a significant technical challenge.

A crucial step in the successful application of MAS in breeding programmes is the development of cheap and easy-to-use molecular markers, such as single-nucleotide polymorphisms (SNPs). Until recently, most wheat laboratories used microsatellite markers in their MAS projects. However, for wheat the new type of SNP marker is becoming more readily available.

This new approach has been adopted by wheat genome scientists at the School of Biological Sciences, University of Bristol, led by Professor Keith Edwards.

"Our aim is to identify tools and technologies to make it easier to do smart breeding to create new varieties of wheat. In particular, this would be used to improve yield and pest and disease resistance, including in commercial lines."

The group has identified that, in order to make wheat pre-breeding more efficient, the development of molecular markers capable of tracking any introduced genomic regions in large numbers of lines is vital. According to Professor Edwards:

"SNP analysis is straight forward and can be automated with relative ease but the task of identifying useful SNPs polymorphisms in wheat is problematic due to its complex and large genome."

Molecular markers must be capable of distinguishing between the relatively large numbers of polymorphisms seen in homoeologous and paralogous genes, compared with the relatively infrequent varietal polymorphisms.Homoeologous/paralogous and varietal SNPs have previously been studied and used in polyploid crops. However, these studies have also shown that distinguishing intervarietal markers from intergenomic polymorphisms is complicated and prone to error. These results, together with the large size of the wheat genome, mean that, despite the global importance of wheat, there are still relatively few validated varietal SNP markers in regular use.

The solution

Developing new strategies to enable UK wheat breeders to breed for yield and crop resistance by MAS underpins the UK's strategy to generate improved wheat varieties.

The development of SNP-based genotyping platforms has lead to an increase in the number of protocols available for analysing genetic variation in numerous species.

Despite the inherent challenge of large-scale genotyping in polyploid species, which carries with it a high instance of homoeologous and paralogous genes, a number of platforms have recently been developed to perform high-density genotyping (large numbers of SNPs, with small numbers of individual plants). These have been successfully employed to genotype wheat.

These technologies can be difficult to optimise and, as such, they have yet to be generally adopted by the wheat community. Lack of adoption leaves few options for wheat breeders and geneticists who wish to carry out medium to low-density genotyping on large or very large numbers of individual plants.

The Cereal Genomics group decided to investigate one of the emerging platforms further at The University of Bristol Genomics facility using a high-throughput genotyping procedure from LGC Biosearch Technologies™ – the Kompetitive Allele-Specific PCR (Polymerase Chain Reaction) (KASP™). They used KASP in a study to investigate the development and validation of a panel of 1114 SNPs in hexaploid bread wheat.

The KASP genotyping assay utilises a unique form of competitive allele-specific PCR that delivers highly accurate, bi-allelic scoring of SNPs and indels at specific loci across a wide range of genomic DNA samples. KASP genotyping chemistry provides great flexibility, with the potential to support both low, mid and high throughput studies, reducing costs through eliminating the need for dual-labelled probes.

Professor Edwards explains:

"We selected LGC's KASP genotyping chemistry because it was the only platform on the market that could overcome our challenges. We had confidence in the technological ability of KASP to increase our quality of data, ensuring the accuracy and efficiency of our findings, in addition to improving the credibility and validity of our studies. We worked closely with LGC to both validate SNPs as a marker system in wheat and transfer the SNP markers, as working assays, to UK wheat breeders."

In the study, researchers undertook to mine the UK wheat germplasm for varietal SNPs. To do this, the group used both the available wheat expressed sequence tags (ESTs) present in the public database and next-generation sequencing (NGS), together with a novel sequence alignment and assembly approach to identify varietal SNPs in lines of interest to European wheat breeders. The group then investigated whether these SNPs could be validated by the KASP Assay.

The group needed a solution that would cancel out the difficulties typically associated with such technologies, creating a simpler and more efficient system for genotyping wheat.

Finally, the researchers made use of the existing Avalon x Cadenza mapping population to examine whether the validated SNPs could be efficiently placed onto the existing linkage groups identified within this doubled haploid population. For further details of the experimental procedure, results and discussion please email techsupport@lgcgroup.com or view the application note.

Results

Using the technology of Biosearch Technologies, the team was able to validate SNPs using KASP on 21 wheat varieties. Using the information generated from KASP, the researchers were able to identify and map 480 SNP markers to 21 linkage groups representing chromosomes, substantially more than normally categorised.

Professor Edwards found that KASP offered the flexibility and accuracy needed to handle the growing quantity and diversity of genotyping data.

"Once the study was completed, the facility functionally validated SNP assays that can be used to identify genetic markers associated with key phenotypes/traits, enabling MAS, all which will greatly improve the yield of wheat."

Since using KASP chemistry, the University of Bristol and the wider industry has been able to improve its practices and methodologies for providing accurate and timely data in the field of wheat genotyping. KASP has accelerated and simplified SNP genotyping for consumers, saving time and reducing costs within the industry.

Professor Edwards further explains,

"KASP increases the rate of SNP-based genotyping whilst decreasing the cost of the individual data points. With the help of KASP, our research also highlights the need to capture all of the information associated with the SNP markers, including the sequence surrounding the SNP being tested. KASP's ability to recognise all of this information enables the SNP to be adapted and used in a variety of genotyping platforms as and when they are developed. This is extremely useful, as the most productive wheat for a particular type of land can now be identified quickly and efficiently."

Industry-wide benefits

As a result of the study, Biosearch Technologies, in conjunction with the University of Bristol, have developed a wheat genotyping panel. The panel contains over 5,000 functionally validated SNP assays that are used by researchers/plant breeders to identify genetic markers associated with key phenotypes/traits, and that enable MAS.

Each assay in the wheat panel has been designed and validated using the accurate and robust PCR-based KASP genotyping chemistry. As a result, individuals can now freely obtain information about SNP markers, including the sequence upon which they are based, primers used for their identification, and haplotypes of common UK varieties. The Biosearch Technologies wheat genotyping panel is a useful tool for breeders, helping them to gain information on genetic relationships among individuals, which is vital for line and hybrid development and is essential for increasing the yield of wheat.

Future

Increasing wheat yields is now one of the top priorities for agricultural research and the sustainable intensification of global agriculture. With the use of KASP, there is now a way to increase the yield of wheat in an efficient and stable manner without compromising on food safety.

Professor Edwards comments,

"This study[1] was the first demonstration of KASP-based chemistry to both genotype wheat varieties and generate a linkage map. The results suggest that KASP-based genotyping is sufficiently sensitive and robust to map individual paralogous genes in the wheat genome."

The Cereal Group at the University of Bristol, in conjunction with Biosearch Technologies, has laid a great foundation in the studies of wheat genotyping, one that the group is looking to exploit more. KASP's efficiencies can be extended to accommodate various genotyping studies, and the genomics centre at the University believe that in the not so distant future it will be possible for wheat breeders to achieve one of their most important goals, to rapidly and effectively genotype thousands of plants with a large and flexible number of markers.

Update

Since the initial study[1] in 2011 significant progress has been made by the team. They have developed and validated a panel of 7,228 SNPs (increased from 1,114) in hexaploid bread wheat, in addition to identifying and mapping 3,629 SNP markers (increased from 480) to 21 linkage groups representing chromosomes.

Reference

1. Plant Biotechnol J. 2011 Dec;9(9):1086-99. doi: 10.1111/j.1467-7652.2011.00628.x. Epub 2011 Jun 1. Transcript-specific, singlenucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (Triticum aestivum L.), Allen AM, Barker GL, Berry ST, Coghill JA, Gwilliam R, Kirby S, Robinson P, Brenchley RC, D'Amore R, McKenzie N, Waite D, Hall A, Bevan M, Hall N, Edwards KJ. School of Biological Sciences, University of Bristol, Bristol, UK.

From left to right: Prof. Keith Edwards, Dr Sacha Allen, Dr Amanda Burridge, Mr Paul Wilkinson and Dr Mark Winfield.

University of Bristol

The University of Bristol has a cluster of research and teaching excellence across the life sciences, from bio-nanotechnology to earth system science; biochemistry to experimental psychology; and individuals to societies. Specifically, the School of Biological Sciences studies everything from wheat genomics to vision ecology and animal behaviour, attracting support from worldwide funders, including BBSRC, NERC, Marie Curie and the EU.

Research activity in the School of Biological Sciences is organised into three research themes: Behaviour, Sensory and Neurobiology; Ecological and Evolutionary Processes; and Plant and Pathogen Biology. Within the Plant and Pathogen Biology department, research is being conducted into cereal genomics, where the aim is to explore and understand cereal genomes with a focus on wheat.

The Cereal Genomics group works closely with the University of Bristol Genomics Facility in developing resources in wheat functional genomics for use by wheat breeders and scientists. This involves wheat genotyping using single-nucleotide polymorphism (SNP) discovery and gene expression analysis using transcriptomics. The group is also investigating the organisation and function of mismatch repair genes with the specific aim of generating new alleles for the wheat breeders. They are also heavily involved in the wheat genome sequencing project.

Established in 2003, The University of Bristol Genomics Facility provides numerous services for Biological Sciences, including; next-generation sequencing, bioinformatics, microarray analysis, capillary sequencing and genotyping, real-time quantitative PCR, and quality assessment of DNA and RNA analysis. One of the University's main areas of research is looking specifically at Cereal Functional Genomics, with the aim of understanding the different varieties of wheat to enhance performance and increase productivity. Ultimately, the institution aims to provide a hub for wheat functional genomics providing UK researchers with authenticated data and biological resources, which are freely available without restriction.

@LGCBiosearch | biosearchtech.com

GEN/0461/LH/0720

BIOSEARCH™ TECHNOLOGIES

GENOMIC ANALYSIS BY LGC

Currently, oil palm plantations produce an average of 21 tonnes of fruit per hectare, which generates 3.74 tonnes of palm oil per hectare, 0.4 tonnes of palm kernel oil and 0.4 tonnes of palm kernel expeller.

Customer case study

Sime Darby Plantation solves the trade-off between trait accuracy and higher throughput costs for oil palm with KASP genotyping chemistry.

Background on oil palm

Palm oil accounts for 38% of the world's vegetable oil market and is harvested from the fruits of the oil palm plant (*Elais guineensis*). Oil palms grow in tropical environments, across the equator and up to 10º north and south[1]. They grow up to 20 metres tall over the average commercial lifespan of 25-30 years. Oil palms have stable commercial fruit production after three years, although they start flowering much earlier. After this time, fruits can be harvested throughout the year. Oil palm fruits grow in bunches that are referred to

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC

in the industry as fresh fruit bunches (FFB), and fruitlets are typically comprised of 45-50% oil.

Currently, oil palm plantations produce an average of 21 tonnes of fruit per hectare, which generates 3.74 tonnes of palm oil per hectare, 0.4 tonnes of palm kernel oil and 0.4 tonnes of palm kernel expeller. Palm oil is extracted from the orange flesh of the fruit known as mesocarp. The fruit kernel or nut (white) is crushed to extract palm kernel oil, leaving behind palm kernel expeller (or cake).

The products extracted from oil palm fruits are used in a wide range of markets including food, personal care and cosmetics (primarily palm kernel oil), biofuel and energy, animal feed (palm kernel expeller) and pharmaceuticals. Palm oil and palm kernel oil-based ingredients are found in approximately 50% of products on our supermarket shelves.

Industry challenge

In the early 19th century, European-run oil palm plantations were first set up in Central Africa and South East Asia. Oil palm was introduced to Malaysia in 1910 by William Sime and Henry Darby. Since the 1940s, improvements to the process for refining of palm oil and the development of enhanced transport networks resulted in the adoption of palm oil for use in Western foods. In this growing market, Malaysia then emerged as the world's largest palm oil producer across the 1970s-1990s, and maintains this position today.

As the global population continues to grow, the demand for food will also increase. Oil palm is the most efficient oil crop available, and is therefore a crucial commodity in feeding the world's expanding population. The oil yield of Malaysian palms has been stagnant at 3-4 tonnes/hectare/year for around 25 years, and increasing palm oil production via increasing the land area of plantations is not sustainable. Sime Darby Plantation is committed to producing certified sustainable palm oil with no deforestation and no peat planting. Palm production also faces new challenges including climate change and agricultural land constraints. It is imperative that the demand for increased palm oil yield is addressed through more sustainable approaches. Other traits of importance to the industry include those that improve ease of harvestability such as long stalks and semi-dwarf palms.

Technical challenge

Plantation owners can use breeding programmes to select for oil palms with desirable traits; however, conventional breeding methods are slow due to the number of years it takes for the oil palms to reach maturity.

Many plantation owners use clonal propagation by tissue culture methods to reproduce their best plants. This method has a high cost and can result in somaclonal variation, potentially causing a loss of key traits. Planting clonal palms also leads to rapid genetic uniformity, which can lead to the crop being vulnerable to new and existing pests and diseases.

More recently, genomic applications such as genomic selection (GS) and marker-assisted selection (MAS) have been used to accelerate breeding programmes, assessing plants for key traits while they are still in the seedling phase.

Marker-assisted selection involves the selection of individuals based on markers within the genome that are linked to traits of interest. This indirect process can be used to select individuals for breeding programmes with a view to enhancing the desired characteristic(s) in subsequent populations. However, classic MAS has its own limitations when oil palm breeders saw a poor selection response when using a limited number of major trait-linked markers.

Markers can also be used to build a predictive model using a genomic selection approach, based on individuals of known phenotype and genotype. This model can be used to rank individuals of unknown phenotype for subsequent selection with greater accuracy. In the breeding of commercial perennial crops, such as oil palm, the number of progeny to assess in each generation is vastly greater than in livestock breeding programmes. As a consequence, the process of genomic selection can be significantly more expensive as vast numbers of samples need to be screened for all markers.

The technical challenge in oil palm breeding is to reduce the required marker density without impacting the phenotype prediction accuracy, thus ensuring that genomic selection is an economically viable approach.

High-throughput processes also need to be implemented to ensure individual plants can be sampled, processed and genotyped accurately, quickly and efficiently. It is crucial that desirable individuals are selected in a timely manner, in order to be transferred into the field at the correct growth stage.

The solution

A high-throughput and easy-to-use genomics technology is required to facilitate detection of genotypes of all young oil palm plants, to enable relevant phenotypes to be accurately predicted without having to raise palms to maturity.

KASP™ genotyping chemistry utilises a unique form of competitive allele-specific PCR that delivers highly accurate bi-allelic scoring of SNPs and indels at specific loci. It is able to support low-, mid-, and high-throughput studies, reducing costs through elimination of the need for dual-labelled probes.

A typical KASP Assay consists of three primers: two allele-specific primers and one common reverse primer. One allele-specific primer is designed to the wild type sequence, and the second allele-specific primer is designed to the sequence of the mutation. Both allele-specifics work in conjunction with the common reverse primer. The KASP Assay is run across a range of DNA samples and the fluorescent signal generated for each sample plotted on a Cartesian (or cluster) plot. The position of each sample on the plot is representative of its genotype. (Watch how KASP works in these videos.) The SNPline™ system from LGC Biosearch Technologies™ facilitates high-throughput processing of samples, from DNA extraction to generation of SNP genotyping data. As the SNPline is modular, it can be adapted to meet the needs of the laboratory, preventing any bottlenecks in sample processing. Components of a SNPline include the repliKator™ for DNA transfer, Meridian™ dispensers for dispensing assays into PCR plates and the Hydrocycler²™ for waterbath thermal cycling.

Results

The GenomeSelect™ programme was developed involving multi-disciplinary teams from within Sime Darby Plantation including breeders, molecular biologists, bioinformaticians, tissue culture specialists, agronomists, IT, and operations. Over 132 genome sequences were analysed, 200,000 SNPs detected, and oil yield traits assessed during the development stage.[3] New laboratories and IT systems were also implemented to facilitate and support the approach.

Sime Darby Plantation used a commercial population of oil palms to investigate a range of predictive models for genomic selection.[2] All individuals were genotyped and phenotyped for a wide range of markers relating to oil yield traits including shell-to-fruit ratio, fruit-per-bunch, and oil-per-bunch, and this information then was used to develop the genomic selection models. All methods resulted in high prediction accuracy, which correlated with the calculated heritability of each of the traits. Further research, including consideration of linkage disequilibrium, which ensures selection of markers remain firmly associated with traits or haplotype blocks in the genome, enabled Sime Darby Plantation researchers to effectively reduce the number of markers for each trait without significantly impacting the accuracy of trait prediction.[3]

By reducing the list of markers required to accurately predict phenotypes of young plants, Sime Darby Plantation was then able to work with Biosearch Technologies to develop a smaller set of KASP Assays to genotype their samples.

A SNPline system was also installed at Sime Darby Plantation to facilitate high-throughput genotyping of thousands of samples. Once the installing process was completed (in June 2015), Sime Darby Plantation was able to routinely extract DNA from 30,000 samples per month, generate 5,000,000 datapoints per month and eventually become one of the biggest SNPline genotyping facilities in Asia. Following extensive analysis of the 80 million generated genotyping datapoints, the first GenomeSelect™ palms were planted in 2016. Sime Darby Plantation aims that by 2023, all replanting will utilise GenomeSelect™ materials.

As a result of the development and implementation of the process for genomic selection in oil palm, in 2017, Sime Darby Plantation's GenomeSelect™ Oil Palm won the prestigious bronze Edison Award for 'Sustainability' as well as 'Best Product Innovation' at the Malaysia Dutch Business Council Innovation & Sustainability Awards.

Industry-wide benefits

The ability to accurately predict the potential phenotype of young plants will be advantageous to the oil palm industry as it will facilitate an increase in the percentage of desirable lines that are planted into the plantations including high yield, drought tolerance and easier harvesting. When these reach maturity, the yield of the plantation is expected to be significantly increased.

Such predictive models enabled by genomic selection drastically reduces the breeding cycle from 20 years by at least half. The development of a straightforward haplotype screening process using KASP chemistry and the SNPline system will enable plantations to quickly and cost-effectively assess the markers of interest in plants while they are still in the nursery stage. This will have a huge impact on the industry as productivity can be sustainably increased, without having to expand area of plantations, resulting in cost-savings and minimal impact on the environment.

Dr. Sukganah Apparow, the Head of Molecular Breeding Laboratory (MBL) at Sime Darby Plantation Technology Centre described the SNPline platform as flexible, robust and user-friendly. Flexibility of both, the SNPline platform and KASP chemistry, enabled Sime Darby Plantation to quickly scale instrumentation and processes necessary to achieve the throughput required for the GenomeSelect™ programme. The programme aims to increase the number of samples while simultaneously decreasing the total number of genetic markers that need to be assessed through haplotype selection. Dr. Apparow added that the SNPline platform is available at Sime Darby Plantation Technology Centre (SDPTC) as a service to screen for high yielding and good quality planting materials. MBL currently provides oil palm DNA testing services for legitimacy, pollen purity, fruit form identification, marker-assisted selection, genetic diversity assessment and others.

Dr. David Ross Appleton, Head of Biotechnology and Breeding at Sime Darby Plantation said,

"The partnership with LGC was a success story, both the in-sourced and outsourced elements of the partnership. LGC not only contributed to develop the in-house DNA extraction and high-throughput KASP genotyping capabilities at SDPTC through extensive training by LGC staff, experienced with operating high-throughput laboratory; but was also able to deliver an outsourced solution to augment capacity using equipment, chemistry and data analysis tools identical to those available at SDPTC."

The work detailed here, performed by Sime Darby Plantation in conjunction with Biosearch Technologies, provides a good reference study for other crop and animal breeding programmes.

The future

Increasing yield is one of the top priorities for agricultural research and sustainable intensification of global agriculture. By utilising KASP chemistry and SNPline instrumentation for haplotype selection, there is now a way to easily and confidently identify higher yielding individuals during the early stages of a plant's development.

Sime Darby Plantation aims to increase the percentage of GenomeSelect™ materials used in replanting each year, eventually achieving 100% in 2023. Selecting for yield across all environments is essential for the future. Many sustainability initiatives that were initially pioneered by Sime Darby Plantation are now considered best practices in the industry, including good water management and the zero burning replanting. The hope is that other plantations will follow the example of Sime Darby Plantation in their approach to increasing yields to meet the growing global demand for palm oil.

Reference

1. Corley RHV and Tinker PB (2003). The Oil Palm. 4th Edition. Oxford: Blackwell Publishing, pp. 3.

2. Teh CK, Ong AL, Kwong QB, Sukganah A, Chew FT, Mayes S, Mohamed M, Appleton D and Kulaveerasingam H. (2016) Genome- wide association study identifies three key loci for high mesocarp oil content in perennial crop oil palm. Sci. Rep. 6: 19075

3. Kwong QB, Ong AL, Teh CK, Chew FT, Tammi MT, Mayes S, Kulaveerasingam H, Yeoh SH, Harikrishna JA, and Appleton D. (2017) Genomic Selection in Commercial Perennial Crops: Applicability and Improvement in Oil Palm (Elaeis guineensis Jacq.) Sci. Rep. 7: 2872.

The GenomeSelect™ Project Team with courtesy of Sime Darby Plantation.

About Sime Darby Plantation

Sime Darby Plantation is the world's largest oil palm plantation company by planted area, accounting for about 4% of total global production of Crude Palm Oil (CPO), with a strong focus on operational excellence, research, innovation and sustainability.

As an integrated plantation company, Sime Darby Plantation is involved in the full spectrum of the palm oil value chain. In Upstream, Sime Darby Plantation operates and manages 241 plantation estates and 72 palm oil mills located in four countries: Malaysia, Indonesia, Papua New Guinea (PNG) and the Solomon Islands, which consist predominantly of oil palm cultivation, harvesting and milling.

Its downstream business, which spans across 14 countries worldwide- Malaysia, Singapore, Indonesia, South Korea, India, Thailand, Vietnam, Japan, China, Germany, United Kingdom, South Africa, Netherlands, United States of America, PNG and Solomon Islands, involves the manufacturing as well as the sales and marketing activities of oils and fats products, oleochemicals, palm oil-based biodiesel, as well as other palm oil derivatives and renewables.

The company also invests heavily in R&D and is the first in the world to successfully sequence, assemble and annotate the oil palm genome. This breakthrough in genome science has resulted in the GenomeSelect Planting Material which was awarded the Edison Award in 2017.

Sime Darby Plantation's Sustainability Report 2016 bagged the Highly Commended Award for Asia's Best Carbon Disclosure at the Asia Sustainability Reporting Awards (ASRA) 2017 in Singapore on 6 February 2018. In line with its commitment to be a leader in sustainability, today, Sime Darby Plantation is also the world's largest producer of Certified Sustainable Palm Oil (CSPO) with an annual production of 2.2 million tonnes. Sime Darby Plantation is also one of the founding members of Roundtable on Sustainable Palm Oil (RSPO), a global multi-stakeholder initiative that ensures companies take ownership in producing sustainable palm oil. Sime Darby Plantation also won the Industry Excellence Award for Plantation category and the Excellence Award for Sustainability Practices at the Minority Shareholder Watchdog Group (MSWG) – Asean Corporate Governance Recognition 2017 for its best practices in sustainability.

As a leader in plantation sustainability, Sime Darby Plantation is committed to its role in the development and promotion of sustainable practices and products in the palm oil sector. Sime Darby Plantation also welcomes partnerships with like-minded organisations to pursue innovative approaches towards developing sustainable products and services in plantation and other related industries.

(L-R) Datuk Franki Anthony Dass and Dr David Ross Appleton. Sime Darby Plantation won the Bronze award for GenomeSelect™ Project in the Sustainability category at the 2017 Edison Awards held in New York on 20 April. Picture with courtesy of Sime Darby Plantation.

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC

Using the genomic sequence of pepper, they examined genes found within the genetic region conferring resistance and found a gene that encodes an enzyme that is highly associated with resistance in pepper.

Customer case study

Seed Biotechnology Center, University of California, Davis - pepper genomics

Pepper – *Capsicum annuum*

Capiscum spp. is an important vegetable and spice crop worldwide and shows incredible diversity in its many varieties.

As with most commercial crops, pepper is the subject of many breeding programmes. These programmes aim to enhance valuable traits in the crop such as improved productivity and resistance to biotic and environmental stresses, as well as consumer characteristics such as taste and colour. For pepper breeders, many of the traits of interest in current *Capsicum* breeding programmes include complex, multi-genic traits that are not easily integrated through traditional breeding strategies.

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC

Phytophthora capsici – a problem for all seasons

One of the biggest problems the pepper industry faces is *Phytophthora capsici* (*Pc*), root rot or late blight, from which most commercial pepper varieties suffer yield losses despite good management practices and resistance strategies.

The lifecycle of *Phytophthora* makes it particularly difficult to get rid off once it infects, and many farmers are forced to abandon their crops if it appears. Selective breeding of varieties, which naturally carry *Phytophthora* resistance, is an attractive long-term solution, but is complicated by the dynamic array of races of *Phytophthora* found in fields over time and variable resistance across plant varieties and tissue types.

Identifying and selectively breeding for pepper strains resistant to *Pc* is exactly what Allen Van Deynze and his fellow researchers at the Seed Biotechnology Center have set about achieving with significant progress, as reported in their recent paper "CaDMR1 cosegregates with QTL Pc5.1 for resistance to *Phytophthora capsici* in Pepper (*Capsicum annuum*)"[1].

The group at the Seed Biotechnology Center is applying the most up-to-date tools and understanding of genetics, genomics and molecular biology to characterise peppers and their diverse characteristics. This, in turn, enables introgression of valuable complex traits, such as disease resistance, into the available breeding stock while keeping other characterisitcs, such as shape, size and taste that are familiar to the consumers.

The goal of the research reported in Rehrig et al.[1] was to identify genes associated with resistance to root rot in pepper. Their approach was to screen a pepper population that exhibited *Pc* resistance and to track common regions in the genome (Quantitative Trait Loci or QTL), and then the genes themselves, that confer the resistance using SNPs (single nucleotide polymorphisms) as genetic markers.

The study identified a gene that encodes an enzyme which is highly associated with *Pc* resistance in pepper; this is the first report to relate candidate genes to a known *Pc* resistance factor in pepper.

High-throughput polymorphism detection and gene expression analysis – the first step

Major progress in pepper breeding was announced by the team in 2013, when the Seed Biotechnology Centre and the Genome Centre published a paper describing a method for high-throughput polymorphism detection and application in the pepper2. Taking advantage of recent *Capsicum* EST sequencing efforts[3] along with a custom

genotyping array design, hybridisation methods and algorithms[4], an Affymetrix Pepper GeneChip® was designed based on > 30,000 pepper genes and the array was then used to identify and analyse polymorphisms across 43 different *Capsicum* lines and *Capsicum* species.

Tracking down genetic determinants of *Pc* resistance

Then, as reported in the study, a high density map with 3892 SNP markers was generated in a set of recombinant inbred lines derived from the highly *Pc* resistant strains (*Capsicum annuum* accession Criollo de Morelos-334 (CM334) and Early Jalapeño). The lines were intensively screened for root rot resistance against *Pc* isolates from across North and South America. QTL effective across, and specific to, isolates were identified and SNP markers which were associated with these resistance QTL were designed and validated across different *Pc*-resistant populations.

Using the genomic sequence of pepper, they examined genes found within the genetic region conferring resistance and found a gene that encodes an enzyme that is highly associated with resistance in pepper.

Now validated as markers of *Pc* resistance, the identified SNPs can be used for targeted breeding and selection of resistant commercial pepper strains with huge potential benefits for both breeder and consumer in the future.

How KASP SNP genotyping was used

In the final stages of this research, QTL and finally a sub-set of SNP markers were identified which acted as genetic markers for *Pc* resistance in multiple pepper strains.

At the point where the team had narrowed their selection of SNP markers down, the array chip was no longer practical and the group switched to using KASP genotyping to follow up on the candidates and to perform the crucial functional validation of them as *Pc* resistance markers.

The research team at the Seed Biotechnology Center used LGC Biosearch Technologies™ to perform KASP genotyping of 30 gene markers in 250 pepper lines, as well as an additional 10 markers in 250 lines out of our US service labs.

The benefit of using KASP SNP markers

The group used the KASP genotyping service, from Biosearch Technologies, to generate the SNP genotype data they required for the functional validation of selected SNP markers.

"LGC's high-throughput and fast turnaround times (40 SNP markers each run on 250 different plant lines in 6 weeks) produced high-resolution data that was crucial for validating our research and getting it published ahead of the field." said Van Deynze.

Seed Biotechnology Center at UC Davis

Allen Van Deynze is director of research at the Seed Biotechnology Center, UC Davis - a leading light for modern plant breeding across both the academic and commercial plant breeding communities. UC Davis is the number one school in the world for agriculture and food studies and is the most cited for scientific publications (QS World University Rankings).

The Seed Biotechnology Center's research team deciphers the genomes of many common crops in order to link plant genes to desirable characteristics, which can lead to improved varieties. Crops such as carrots, cotton, lettuce, melons, peppers and spinach are investigated and improved, often in partnership with key breeders of these crops.

References

1. Rehrig, W. Z., Ashrafi, H., Hill, T., Prince, J., & Van Deynze, A. (2014). CaDMR1 Cosegregates with QTL Pc5.1 for Resistance to *Phytophthora capsici* in Pepper (*Capsicum annuum*). The Plant Genome, 7(2).

2. Hill, T. A., Ashrafi, H., Reyes-Chin-Wo, S., Yao, J., Stoffel, K., Truco, M. J., ... & Van Deynze, A. (2013). Characterization of *Capsicum annuum* genetic diversity and population structure based on parallel polymorphism discovery with a 30K unigene pepper GeneChip. PloS one, 8(2), e56200.

3. Kim, H. J., Baek, K. H., Lee, S. W., Kim, J., Lee, B. W., Cho, H. S., ... & Hur, C. G. (2008). Pepper EST database: comprehensive in silico tool for analyzing the chili pepper (*Capsicum annuum*) transcriptome. BMC plant biology, 8(1), 101.

4. Stoffel, K., van Leeuwen, H., Kozik, A., Caldwell, D., Ashrafi, H., Cui, X., ... & Van Deynze, A. (2012). Development and application of a 6.5 million feature Affymetrix Genechip® for massively parallel discovery of single position polymorphisms in lettuce (*Lactuca spp.*). BMC genomics, 13(1)

@LGCBiosearch | biosearchtech.com

GEN/807/LH/0221

BIOSEARCH™ TECHNOLOGIES

GENOMIC ANALYSIS BY LGC

Application note

Variation in the ovocalyxin – 32 gene in commercial egg-laying chickens and its relationship with egg production and egg quality traits

J. E. Fulton, Hy-Line International, USA; LGC Biosearch Technologies, Hoddesdon, UK.

Summary

The quality of eggshells is an important trait for commercial egg production. A number of studies have linked polymorphisms in the gene encoding the eggshell protein ovocalyxin-32 (OCX32) with eggshell features including strength and thickness, indicating that variants in the gene may be relevant to the selection of commercial egg-laying poultry lines.

In the study of Fulton[1], exons 2-6 of the OCX32 gene were sequenced in multiple elite commercial egg-laying lines, and SNP detection and analysis carried out using KASP™ genotyping chemistry from Biosearch Technologies™ to identify multiple polymorphisms. The genotype data was used to identify changes in amino acids, infer novel protein haplotypes, and associate these protein variations with a range of egg traits.

The study identified 28 SNPs and 1 SNP/indel in exons 2-6 of the OCX32 gene, which encompasses 78% of the gene coding region. SNP analysis data indicated that the poultry lines tested carry 19 different variants of the OCX32 protein. Trait association studies indicated that the variants were linked to different degrees with traits including eggshell colour, early egg weight, albumen height, puncture score and yolk weight. Selection pressure for some variants over time was also evident in three of the poultry lines, indicating that some polymorphisms in the OCX32 gene may confer changes in egg traits that are desirable for commercial egg production.

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC

Application note

Variation in the ovocalyxin – 32 gene in commercial egg-laying chickens and its relationship with egg production and egg quality traits

Introduction

Eggshell structure and strength are important traits for commercial egg production, as they can determine whether an egg will be able to withstand handling and transportation, as well as microbial challenge. Numerous studies in the literature have reported relationships between single nucleotide polymorphisms (SNPs) in the ovacalyxin-32 gene and egg-related traits, including eggshell thickness weight and stiffness, in a number of commercial poultry lines.

OCX32 is a 32 kDa matrix protein that is expressed in the avian uterus and isthmus, and incorporated as a component of the outer layers of the eggshell and the shell cuticle. In this study, the researchers used KASP genotyping chemistry from Biosearch Technologies to identify and analyse SNPs spanning exons 2-6 of the OCX32 gene in eight elite commercial brown and white egg-laying poultry lines, determine amino acid alterations in the protien, and infer exon and protein haplotypes within individual poultry lines.

Trait association studies demonstrated significant effects of OCX32 polymorphisms on a number of egg characteristics, providing an insight into the selection pressures for certain variants of the OCX32 gene.

Trait association studies demonstrated significant effects of OCX32 polymorphisms on a number of egg characteristics, providing an insight into the selection pressures for certain variants of the OCX32.

Application note

Variation in the ovocalyxin – 32 gene in commercial egg-laying chickens and its relationship with egg production and egg quality traits

Materials and methods

Sequencing

The study focused on eight elite brown and white eggshell commercial egg-laying lines from three different breeds:

- White Leghorn (white eggshells)
- White Plymouth Rock-derived lines (brown eggshells)
- Rhode Island Red (brown eggshells).

Purified amplicons (exons 2-6) were sequenced by SeqWright, and sequence electopherograms analysed by sequencher 4.9 (Gene Codes Corp.). Sequences were aligned with the Jungle Fowl (JF) genome sequences accessed from UCSC.

SNP detection and genotyping

Initially, SNPs were detected by PCR amplification to yield different products depending on the genotype, followed by gel-based detection. SNP detection was subsequently changed to a more rapid and less expensive method using KASP chemistry, a competitive allele-specific PCR-based fluorescent SNP genotyping system.

Traits

Both male and female traits were evaluated. Egg production and quality traits assessed included egg weight, shell colour, egg production, albumen height, yolk weight and body weight, during early production and, where possible, during late production.

Statistical analysis

Power-Marker software[2] was used for haplotying individuals, for LD analysis to calculate the frequency of haplotypes, and for evaluating phenotypic effects of protein haplotypes (PHTs). Association of PHTs with traits associated with egg production and egg quality was evaluated using a haplotype trend regression (HTR) option of Power-Marker[3].

Results

DNA polymorphisms

Exons 2-6 represent 78% of the OCX32 cDNA, cover 588 bases of exon sequence, and encode 196 amino acids of the OCX32 protein. Within the gene region analysed, 28 SNPs and one SNP/InDel (insertion/deletion) were identified, 15 of which had not previously been described. All but three of the exonic SNPs resulted in a predicted change in protein sequence, and most of the polymorphic sites were found in multiple lines, which suggests that they represent a major fraction of the common polymorphic sites found in commercial white and brown egg lines. Four of the SNPs resulted in amino acid substitutions that may alter OCX32 protein 3D structure or function.

Exon haplotypes

The researchers generated a minimum panel of eight SNPs for genotyping each poultry line. Data from these eight SNPs was used to identify exon amino acid haplotypes in large numbers of individuals from each line. Because all of the SNPs used to identify haplotypes resulted in amino acid changes, the SNPs effectively represent different protein haplotypes.

From this data it was found that exon 2 of the OCX32 gene can present as one of three exon haplotypes, including one that exhibits six SNP variants that always occur together. Exons 3 and 5 each harboured a single SNP, and present two haplotypes each.
Exon 4 demonstrated two independent SNPs, and thus four haplotypes, while exon 6 also exhibited four haplotypes, one of which has five SNP variants that always present together.

Application note

Variation in the ovocalyxin – 32 gene in commercial egg-laying chickens and its relationship with egg production and egg quality traits

The combination of SNPs present in the OCX32 gene resulted in 19 different proteins possible across the eight lines. The keys here is that the approach used in this uncovered far greater diversity than could be identified using single SNP analysis, even within intensively selected lines.

Protein variation and egg traits

Calculated allele frequencies for 13 SNPs, averaged across all generations of both males and females, were additionally evaluated in terms of egg traits.

The most common effects were seen for shell colour in all five white egg lines, but not in the brown eggs. Significant effects were found for albumen height in a number of brown and white egg lines. Associations were also found for early and late egg weight in five lines, and there was some support for an association with puncture score in three lines, indicating that OCX32 may have an impact on these traits as well. In three of the poultry lines there was a substantial change in the level of variation in the OCX32 gene and its protein among selectively bred poultry, indicating either that there may be a selective advantage for some variants, or perhaps variation itself is advantageous. Further evidence for this was indicated by the finding that there were significant changes in the frequency of some haplotypes in three of the poultry lines over generations, with some variants increasing in frequency, and others decreasing.

Conclusions

The results demonstrate a large potential for variation within the OCX32 gene, even among a relatively small subset of selectively bred poultry lines. In this study 86% of the variants detected resulted in amino acid changes, with the identification of 19 different OCX32 protein variants within the eight commercial poultry lines tested. Further evaluation of the effects of variation in the OCX32 gene, and the effects of these variants on egg-production and egg structure or quality traits could potentially help direct future selection strategies for commercial egg-laying poultry lines.

Use of the KASP chemistry for genotyping has facilitated the evaluation of genetic variation across five exons of a gene that has previously been implicated in egg quality. In contrast with other studies, this work identified a large number of SNPs not previously reported, in a far greater number of poultry lines, and thus encompassing a much larger gene pool. The data showed that multiple, complex OCX32 protein haplotypes are found in commercial egg-laying poultry lines that have been selected through intensive breeding.

Key words: eggshell color, eggshell quality, Ovocalyxin-32, protein haplotypes

References

1. Fulton, J.E., Soller, M., Lund, A.R., Arango, J. and Lipkin, E. (2012) Variation in the ovocalyxin-32 gene in commercial egg-laying chickens and its relationship with egg production and egg quality traits. Animal Genetics 43, s1, 102-113.

2. Liu, K. and Muse, S.V. (2005) POWERMAKER: an integrated analysis environment for genetic marker analysis. Bioinformatics 21, 2128-9.

3. Dmitri, V., Zaykin, P.H., Westfall, S., Young, S., Karnoub, M.A., Wagner, M.J. and Ehm, M.G. (2002) Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. Human Heredity 53, 79-91.

@LGCBiosearch | biosearchtech.com

GEN/810/LH/0421

BIOSEARCH™ TECHNOLOGIES

GENOMIC ANALYSIS BY LGC

Today's laying hens are each capable of producing over 300 eggs per year. The world's human population is growing rapidly and the high-quality protein and nutrient-rich egg is an inexpensive and transportable source of nutrition for this growing population.

Customer case study

Hy-Line International - poultry genomics

Why poultry?

Between 2000 and 2010, the global egg output expanded by more than 25% from 51 million tonnes to 63.8 million tonnes. World egg production will likely reach a record 65.5 million tonnes in 2013[1].

In order to increase the ability of the hens to produce these eggs, and assure that these eggs are of the highest quality, research regarding their genetic make-up is required to make sure the best individuals are identified and used to breed the next generation to meet the growing production requirements.

Poultry breeding uses a combination of experience, expertise and science that requires long term investments of both time and resources. Certain traits (genetically determined characteristics) such as feed consumption, egg shell quality and rate of lay can be measured and analysed by geneticists involved with poultry breeding.

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC

Each bird is evaluated for the full range of useful characteristics and compared against others. Elite birds which exhibit these characteristics of better performance are then used to breed and produce the next generation of egg-laying hens.

These superior birds have naturally occurring changes (variation) within their DNA that impacts their performance. One type of DNA sequence variation commonly used for this type of analysis is single nucleotide polymorphisms or SNPs. Recent research has identified over 39 million novel variations that exist within the DNA code of chickens.

Detection of SNPs by genotyping is an important tool for identifying variation in important trait genes and understanding how this variation can affect the traits. Such techniques can help poultry breeding research scientists to identify and measure the genetic variation of traits which are important to commercial egg production and bird welfare.

Industry challenge

It is estimated that there are between 5 and 5.5 billion egg-laying hens in the world.

Today's laying hens are each capable of producing over 300 eggs per year. The world's human population is growing rapidly and the high-quality protein and nutrient-rich egg is an inexpensive and transportable source of nutrition for this growing population.

The challenge for the poultry industry is to produce this food source in an efficient way. A decreased input is needed to produce each egg, to ensure that each egg can reach the consumer while at the same time decreasing the carbon footprint of the laying hen.

Hy-Line International strives to create the highest standard of commercial egg-laying hen with good liveability, feed efficiency, nesting behaviour and persistency of lay that also produces a good quality shell and subsequently the best egg to get to market.

Hy-Line International uses genetic research and testing to combat these challenges. By identifying variation within the DNA of the egg-laying bird, they can decrease the amount of feed needed to produce each egg. Variation within the DNA that affects the proteins in the egg shell can be identified to improve the quality of the shells so that more eggs can reach the consumer. The carbon footprint caused by producing these eggs can also be addressed indirectly by improving production parameters.

Technical challenge

The Hy-Line International research programme has been operating since 1936. Within the programme, Hy-Line measures over 30 different traits of importance to commercial egg production and bird welfare. Information from thousands of birds within both the research farms and the commercial field test programmes is fed into extensive databases and subjected to extensive statistical analyses. The challenge is to integrate information from molecular genetics into the existing breeding programme system.

The company has developed the world's most extensive DNA archive in the poultry industry consisting of samples from over 15 generations of birds from multiple lines. Each of these samples is associated with trait information as well as information on all family members. The Hy-Line breeding programme involves identification of multiple SNPs within genetic regions of interest and the subsequent genotyping of large number of individuals within multiple lines. Application of this information into the breeding programme requires that the results be provided rapidly to enable selection of the next generation of breeding birds. Dr Janet Fulton, Molecular Geneticist at Hy-Line International comments:

"When the Molecular Genetics programme was initiated in 1996, the state-of-the-art technology at the time allowed us to study 100 different DNA variations (or markers). Bird-to-bird variation was identified in 5,000 samples per week. However, the requirements of our current research meant that we needed to expand the number of samples tested without expanding the costs. With the increased numbers of samples we also needed to ensure that every sample and every result was correctly identified and tracked."

The scientists at Hy-Line sought a technology that could analyse hundreds of thousands of samples quickly and that could do so without large financial resources or disruption to their current analysis procedure. The scientists depend upon the results in order to complete the breeding cycle. This means that very accurate data must be delivered within a very short time.

In a recent study[2], researchers at Hy-Line identified *ovocalyxin-32* (*OCX32*), as a candidate gene for selection for eggshell traits in commercial poultry populations. The OCX32 protein is a component of the avian uterine milieu present during formation of the eggshell. Although the specific function of *OCX32* in eggshell formation and structure is unknown, the presence of the protein as a component of the outer layer of the eggshell suggests it has an important role within the avian eggshell. SNP variation in *OCX32* gene has been reported in a number of poultry breeds and researchers at Hy-Line wanted to identify polymorphic sites within the *OCX32* gene in their breeding stock. Initially, SNP variation was detected using allele-specific PCR primers that result in a different amplicon size for each SNP followed by gel based detection. However, this method was costly, labour intensive and prone to sample tracking errors.

The solution

When identifying variations in the OCX32 gene, SNP detection was changed to a more rapid and less expensive method based on KASP™ genotyping chemistry from LGC, Biosearch Technologies. Hy-Line implemented a high-throughput genotyping method using KASP assays to help detect SNP variations. Hy-Line found that KASP genotyping allowed them to do rapid in-house detection of SNP variations and that the technology permitted them to carry out individual testing of thousands of samples, for hundreds of carefully selected SNPs, at a rate of over 100,000 tests per day.

"The KASP technology provided the best and most rapid identification of genetic variation within our populations."

Dr Fulton adds,
"The volumes and complexity of data generated by this technology is unimaginable. The dedicated Kraken™ software system is used to analyse and track the SNP data and the results are then rapidly integrated into the existing databases that contain the thousands of collected data points gathered from the research farms and global field tests."

Utilising newly developed statistical methods, KASP is used to associate SNP variations with performance. Both the speed and accuracy of the selection process is improved with genomic selection made possible with KASP genotyping. Not only will the right birds now be more accurately identified, but their genetic potential is identified earlier in the birds' life.

Results

KASP genotyping perfectly fits the requirements of very large numbers of samples to be tested over a large number of individual SNP assays; the *OCX32* gene is a perfect example of the use of KASP genotyping. Janet Fulton commends KASP's versatility and accuracy, explaining that:

"You have so much control over what you can do in terms of designing the primers to detect your specific variants. You can detect this variation with different equipment but using the

same materials and you'll get the same results. This makes it easy to adjust assay volumes when moving between low and high numbers of samples and manual to semi-automated systems."

Why KASP?

The KASP genotyping assay utilises a unique form of competitive allele-specific PCR that delivers high levels of assay and accuracy, whilst ensuring cost savings. The technology enables highly accurate bi-allelic scoring of SNPs and indels at specific loci across a wide range of genomic DNA samples. The benefit for Hy-Line's research was KASP's ability for rapid SNP testing while still producing highly accurate results needed for input into a commercial breeding programme.

Future

Hy-Line's research will affect the egg industry on a larger scale than simply improving the quality of egg shells. With this genetic programme, the key performance traits needed to be a successful layer are identified and improved across multiple lines. The genome variation which influences the quality of the shell are found and improved across generations.
Hy-Line's research continues to improve hens through their laying rate, their feed efficiency and their liveability. KASP genotyping facilitates the selection process for those birds that will produce the next generation of elite breeding stock.

References

1. The Poultry Site, Jan 2013, http://www.thepoultrysite.com/articles/2653/global-poultry-trends-world-egg-production-sets-a-record-despite-slower-growth.

2. Fulton, J. E., Soller, M., Lund, A. R., Arango, J., & Lipkin, E. (2012). Variation in the ovocalyxin - 32 gene in commercial egg - laying chickens and its relationship with egg production and egg quality traits. Animal genetics, 43(s1), 102-113

Hy-Line International

Hy-Line International is a world leader in poultry genetics with a rich history of innovation. Founded in 1936, by Henry A. Wallace, Hy-Line International was the first of the modern egg layer (hen) genetics companies to incorporate hybridisation and the potential of hybrid vigor into its breeding programme on a commercial scale. The programme was used alongside time tested methods of genetic selection coupled with scientific statistical analysis to develop and improve one of the world's most extensive gene pools. Today, Hy-Line International continues to be a pioneer as the first company with its own in-house molecular genetics team. It is leading the industry in the application of DNA-based technology to its breeding and genetics programme.

The company has several commercial products across the global marketplace and access to one of the world's most extensive gene pools. Hy-Line International produces and sells both brown and white egg stock to more than 120 countries worldwide and breeds hens known for their superior liveability, feed efficiency and egg quality.

Research is the key component in the genetic development process and Hy-Line geneticists use exacting research procedures to preserve the unique genetic balance while making continuous improvements to the performance of all Hy-Line stock. Hy-Line uses only established and universally accepted methods to detect naturally occurring genetic variation, for evaluation and subsequent selection in their birds. There is no use of genetic modification or cloning anywhere within its breeding programme.

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC

Application note

Trans-regulation of mouse meiotic recombination hotspots by *Rcr1*

PLoS Biol. 2009 Feb 17; 7(2): e36[1]; Petko M. Petkov, The Jackson Laboratory, Bar Harbor, USA; LGC Biosearch Technologies™, Hoddesdon, UK

Summary

Meiotic homologous recombination generates genetic variety amongst offspring and ensures that the chromosomes are segregated accurately during meiosis. Recombination occurs in hotspots, and the aim of this study was to identify *trans*-acting factors that control hotspot positioning in mammals, specifically inbred mice.

The study compared the location of crossovers in an 8 Mb segment of a 100 Mb region of mouse chromosome 1 where the longer region was heterozygous C57BL/6J (B6) x CAST/EiJ (CAST) and the remainder of the genome was either similarly heterozygous or entirely homozygous B6. This lack of CAST alleles had a major effect in hotspot activity, where several hotspots lost recombination activity, additional new hotspots were discovered, and some hotspots were unaffected. This suggested the impact of one or more *trans*-acting genes with CAST alleles that activated or suppressed hotspot activity.

Analysis of activity of three activated hotspots in sperm samples from males resulting from two genetic crosses identified a single *trans*-acting regulator of hotspot activity (*Rcr1*), on chromosome 17 and results from an *Escherichia coli* (*E. coli*) cloning assay suggested that *Rcr1* controlled the formation of the double-strand breaks that initiate the recombination process.

BIOSEARCH™ TECHNOLOGIES
GENOMIC ANALYSIS BY LGC

Application note

Trans-regulation of mouse meiotic recombination hotspots by *Rcr1*

Introduction

Homologous recombination during meiosis is initiated by the formation of DNA double-strand breaks (DSBs) created by the highly conserved topoisomerase IV– like protein SPO11. These DSBs act as the sites for chiasmata and crossovers to form. The DSB repairs can either form a crossover between homologous chromatids (CO; exchange of genetic information) or a non-crossover (NCO; acquisition of genetic information by the initiating chromatid from its partner)[2,3].

In yeast and mammals the recombination rates vary along the chromosome[4,5,6,7,8], and most occur in hotspots[9,10]. Little is currently known about the causes behind the locations and activity of these recombination hotspots in mammals. *Trans*-acting factors controlling hotspot activation have been identified in several cases in yeast.

The aim of the study was to identify any *trans*-acting factors that control hotspot positioning in inbred mice. It focused on crossovers in an 8-Mb segment of a 100 Mb region of mouse chromosome 1, where the longer region was heterozygous C57BL/6J (B6) x CAST/EiJ (CAST) and the remainder of the genome was either similarly heterozygous or entirely homozygous B6.

Meiotic homologous recombination generates genetic variety amongst offspring and ensures that the chromosomes are segregated accurately during meiosis.

Application note

Trans-regulation of mouse meiotic recombination hotspots by *Rcr1*

Materials and methods

The mice

The mice were genetic crosses involving the B6 and CAST mouse strains. In the first cross (interstrain cross), B6 mice were mated to CAST, and the F1 hybrids were backcrossed to B6. In the second cross (congenic cross), B6 mice were mated to B6. CAST-1T, a congenic strain carrying 100 Mb of CAST DNA sequences from distal chromosome 1 introgressed into C57BL/6J.

The resulting F1 hybrids were then backcrossed to B6. The CAST alleles were present in the interstrain B6xCAST F1 animals and absent in the congenic B6xB6. CAST-1T F1 mice, which are homozygous B6 outside the 100-Mb chromosome 1 region.

E. coli cloning assay

The researchers applied the *E. coli* cloning assay to hotspots Hlx1 and Esrrg-1 as described in Ng *et al.*[11]. The hotspot sequence was amplified with primers common to both fragments, and the cloning resulted in each colony representing a single DNA strand from the original meiotic event. An aliquot from each culture underwent fluorescent SNP genotyping.

Phenotyping

The assays used partially purified DNA as described in Paigen *et al.*[8]. To detect recombination activity, recombinant DNA fragments were selectively amplified at the hotspot sequence of interest.

Sperm DNA samples from 12 week old animals underwent two rounds of nested PCR using allele-specific primers in each of the two rounds. These were specific to the B6 alleles flanking the hotspot, and to the CAST alleles. The PCR conditions ensured that the product was only from recombinant and not parental DNA.

Genotyping

The assays used partially purified DNA as described in Paigen *et al.*[8]. Recombination activity in the region of 183.5-191.5 Mb on mouse chromosome 1 was fine-mapped using SNP markers.

For genome-wide association mapping (GWAS), the researchers genotyped all progeny at 20 Mb resolution using the KASP™ genotyping system, with markers selected from The Jackson Laboratory genotyping panel[12].

Fine-mapping of CAST alleles on chromosome 17 used a combination of the microsatellite markers D17Mit48, D17Mit57, D17Mit113, and D17Mit46 and SNP markers.

Results and discussion

Trans-activation and suppression of hotspots

The B6xCAST recombination map was based on 6028 meioses in F1 animals. Among the F1 offspring, 735 had a single CO event in the 8-Mb region of interest (264 in females; 471 in males), which produced a sex-averaged map of 12.2 cM (8.8 cM in females; 15.6 cM in males).

The B6xB6.CAST-1T map was based on 2083 meioses, with 175 leading to single CO events (83 in females; 92 in males), which produced a sex-averaged map of 8.4 cM (7.1 cM in females; 10.1 cM in males).

The COs were mapped to hotspot-level resolution, and while most showed similar activity in the two crosses, six hotspots (*Fbxo28, Dusp10, Hlx1, D1Pas1, Esrrg-1*, and *Kcnk2*) disappeared in the congenic cross and three new and previously undetected hotspots appeared (*Capn2, Kctd3*, and *Ptpn14*). There were statistically significant sex differences in three hotspots in the interstrain cross (*Hlx1, Esrrg-1*, and *Kcnk2*) and one in the congenic cross (*Kctd3*).

Application note

Trans-regulation of mouse meiotic recombination hotspots by *Rcr1*

These results show that the recombination activity of a number of hotspots on chromosome 1 was affected (either activated or suppressed) by CAST allele(s) of distant *trans*-acting loci. Other hotspots were unaffected.

Mapping *Rcr1*, the *trans*-acting locus

The researchers used two mapping crosses to search for *trans*-acting genes regulating recombination at specific hotspots. The first was created by crossing the B6.CAST-1T congenic strain with CAST, and then crossing the F1 females with B6 (Figure 1A), creating mice that were heterozygous B6/CAST at the 100 Mb congenic region, with segregated CAST alleles in the rest of the genome. This allowed the detection of any relevant X-linked genes, meant that all male progeny were informative; and that any CAST alleles were heterozygous.

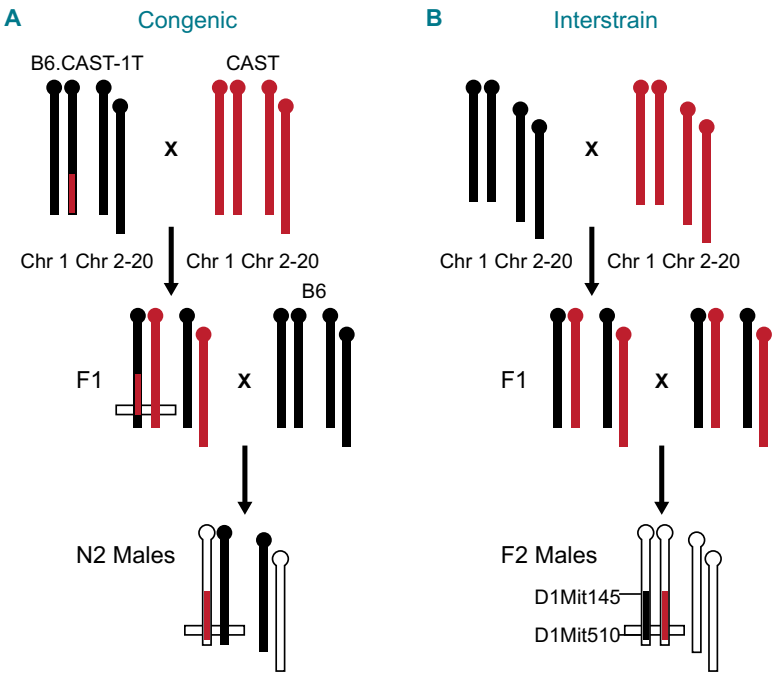


Figure 1. Crosses for mapping of *trans*-acting genes (B6.CAST-1TxCAST)xB6 (A) and B6xCAST (B) crosses used for mapping trans-acting genes. B6 sequences are in black, and CAST in red. The region where the recombination is studied is boxed. Segregating loci are located in the white regions. Note that any recombination within the congenic region between the B6.CAST-1T and CAST chromosomes does not alter the allelic composition of this region.doi:10.1371/journal.pbio.1000036.g003.

From this cross, 211 male animals were phenotyped for activity of the three hotspots *Hlx1*, *Esrrg-1* and *Kcnk2* using allele-specific sperm DNA assays. F2 animals from a second cross, mating B6xCAST F1 mice together, were also phenotyped.

Animals from both crosses were genotyped with 165 SNP markers, and this showed a strong link between hotspot activity and a locus located within a 5.30 Mb window on proximal chromosome 17. Both crosses produced identical map locations, and no other chromosome location showed significant linkage in either cross. This locus was designated Recombination regulator 1 (*Rcr1*).

Application note

Trans-regulation of mouse meiotic recombination hotspots by *Rcr1*

The molecular identity of *Rcr1* is unknown, but the closest known phenotypic parallel is the *ADE6-M26* hotspot in Saccharomyces pombe activated via chromatin remodelling mediated by the *ATF1.PCR1* transcription factor.

Rcr1 controls the initiation steps of the recombination process

If *Rcr1* controls the early stage of recombination, it would control the appearance of both CO and NCO gene conversions at susceptible hotpots. If it acts on the choice between CO and NCO, NCOs would persist at susceptible hotspots in the absence of the *Rcr1* CAST allele.

The researchers carried out a cloning assay counting the number of COs and NCOs at individual hotspots in F1 sperm DNA, and found that both NCOs and COs were absent from these hotspots, suggesting that *Rcr1*'s effect is at the initiation of recombination, before the choice between CO and NCO.

Conclusion

Trans-acting factors have potential to explain variations in hotspot activity, sex differences in the activities of individual hotspots, and the failure to find a consensus DNA sequence that accounts for the specificity of SPO11 cleavage.

In a subsequent study[13], the researchers were able to identify *Prdm9* as a mammalian protein regulating meiotic recombination hotspots which initiates studies of an important biological control system that has hitherto been inaccessible. The results of this successive study show that these sequences represent recombination hotspots. Additional studies in this and several other labs have shown that *Prdm9* controls activation of all, or nearly all, recombination hotspots.

Keywords: PCR, genotyping, fluorescence, allele-specific, recombination, meiosis, *trans*-regulation, chromosomes, crossovers.

References

1. *Trans*-regulation of mouse meiotic recombination hotspots by *Rcr1*. Parvanov, E D, *et al.* Feb 2009, PLoS Biol, Vol. 7.
2. Distribution of meiotic recombination sites. de Massy, B. Sep 2003, Trends Genet, Vol. 19, pp. 514-522.
3. Regulating double-stranded DNA break repair towards crossover or non-crossover during mammalian meiosis. Baudat, F and de Massy, B. 2007, Chromosome Res, Vol. 15, pp. 565-577.
4. Pulsed-field gel analysis of the pattern of DNA double-strand breaks in the Saccharomyces genome during meiosis. Game, J C. 1992, Dev Genet, Vol. 13, pp. 485-497.
5. Where the crossovers are: recombination distributions in mammals. Kauppi, L, Jeffreys, A J and Keeney, S. Jun 2004, Nat Rev Genet, Vol. 5, pp. 413-424.
6. Global mapping of meiotic recombination hotspots and coldspots in the yeast Saccharomyces cerevisiae. Gerton, J L, *et al.* Oct 2000, Proc Natl Acad Sci U S A, Vol. 97, pp. 11383-11390.
7. A fine-scale map of recombination rates and hotspots across the human genome. Myers, S, *et al.* Oct 2005, Science, Vol. 310, pp. 321-324.
8. The recombinational anatomy of a mouse chromosome. Paigen, K, *et al.* Jul 2008, PLoS Genet, Vol. 4.
9. An initiation site for meiotic crossing-over and gene conversion in the mouse. Guillon, H and de Massy, B. Oct 2002, Nat Genet, Vol. 32, pp. 296-299.
10. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. Jeffreys, A J, Kauppi, L and Neumann, R. Oct 2001, Nat Genet, Vol. 29, pp. 217-222.
11. A quantitative assay for crossover and noncrossover molecular events at individual recombination hotspots in both male and female gametes. Ng, S H, *et al.* Oct 2008, Genomics, Vol. 92, pp. 204-209.
12. An efficient SNP system for mouse genome scanning and elucidating strain relationships. Petkov, P M, *et al.* Sep 2004, Genome Res, Vol. 14, pp. 1806-1811.
13. *Prdm9* controls activation of mammalian recombination hotspots. Parvanov ED, Petkov P M, Paigen K. 2010. Science 327:835.

@LGCBiosearch | biosearchtech.com

GEN/808/LH/0621

BIOSEARCH™ TECHNOLOGIES

GENOMIC ANALYSIS BY LGC

Summary

As a changing climate and growing population challenge the world's food supply, innovative teams of scientists are turning to advanced genomics tools to build solutions. And they're partnering with Biosearch Technologies to ensure that these mission critical projects are a success. We hope the genotyping application notes and case studies presented here inspire your efforts to address global problems and demonstrate how the market-leading PCR-based KASP genotyping chemistry can enable and accelerate achievement of your project goals.

Whatever challenge your team is working on, whether the goal is to enhance food quality, ensure food security, or improve productivity with environmentally sustainable practices, Biosearch Technologies has an integrated portfolio of accurate, reliable and high-quality genomic tools that work with you to advance your science and overcome the productivity challenge facing global agriculture. Find out how our mission critical tools are helping our customers innovative in agrigenomics online: [biosearchtech.com/sectors/agrigenomics.](biosearchtech.com/sectors/agrigenomics)

A helping hand from LGC Biosearch Technologies

Sample preparation:

Our range of products including QuickExtract™, Kleargene™ and sbeadex™, allows delivery of optimised extractions and purifications for plant material including leaves, seeds, and other plant matrices.

Oligos and reagents:

Expertise in genotyping and gene expression through PCR products including KASP (KBD/KOD), BHQplex™ CoPrimers™, BHQplus™, BHQ™ Probe Master Mix and enzymes enables discovery to production-scale testing. Assay development and validation options available.

Lab services:

NGS services including GBS, Flex-Seq (targeted genotyping by sequencing) and more. Additional service solutions available ranging from DNA extraction to genotyping to sequencing and bioinformatics.

For Research Use Only. Not for use in diagnostic procedures.

@LGCBiosearch | biosearchtech.com

BIOSEARCH™ TECHNOLOGIES

GENOMIC ANALYSIS BY LGC