

# FINAL REPORT

## Integrating crop improvement technologies for rapid genetic advance

### DAQ00076

#### Project Details

- **Project Code:** DAQ00076
- **Project Title:** Integrating crop improvement technologies for rapid genetic advance
- **Start Date:** 01.10.2004 **End Date:** 09.10.2008
- **Supervisor:** David Jordan (Principal Plant Breeder)
- **Organisation:** Department of Primary Industries and Fisheries  
604 Yangan Road Warwick QLD 4370
- **Contact Name:** David Jordan  
Phone: 07 4660 3622  
Email: David.R.Jordan@deedi.qld.gov.au

#### Summary

Crop improvement practice has not changed fundamentally in the past 50 years and is yet to capture revolutionary changes in information and biological technologies in the integrated manner required for their application to complex traits. The objective of this project was to explore methods to integrate new marker technologies, simulation modelling and statistics with the aim of increasing the rate of genetic gain in breeding programs focused on complex traits such as grain yield. The project was successful in identifying promising opportunities and demonstrating their value in the context of an applied breeding program.

#### Report Disclaimer

This document has been prepared in good faith on the basis of information available at the date of publication without any independent verification. Grains Research & Development Corporation (GRDC) does not guarantee or warrant the accuracy, reliability, completeness or currency of the information in this publication nor its usefulness in achieving any purpose. Readers are responsible for assessing the relevance and accuracy of the content of this publication. GRDC will not be liable for any loss, damage, cost or expense incurred or arising by reason of any person using or relying on information in this publication. Products may be identified by proprietary or trade names to help readers identify particular types of products but this is not, and is not intended to be, an endorsement or recommendation of any product or manufacturer referred to. Other products may perform as well or better than those specifically referred to. Check [www.apvma.gov.au](http://www.apvma.gov.au) and select product registrations listed in PUBCRIS for current information relating to product registration.

#### Copyright

Grains Research and Development Corporation. This publication is copyright. Apart from any use as permitted under the Copyright Act 1968, no part may be reproduced in any form without written permission from the GRDC.

#### Old Reports

The subject matter in this report may have been revisited or may have been wholly or partially superseded in subsequent work funded by GRDC or others (check completion date).

## Conclusions

The objective of the DAQ00076 project was to explore ways of using new technologies to improve the efficiency of crop improvement programs and improve progress for complex traits. The technologies investigated included marker assisted breeding, simulation modelling, increased computer processing speed, informatics and enhanced statistical analysis. One of the main drivers for this research was the idea of adding value to increasingly expensive field based phenotyping for yield using technologies that are either low cost or are declining in cost with the aim being to increase genetic gain per dollar. To this end the main areas of research were 1) developing the capacity to detect quantitative trait loci (QTL) in breeding populations and 2) using simulation modelling to enhance plant breeders' understanding of complex traits and genotype x environment (GxE). Both of these technologies are becoming less expensive and can be used to extract additional value out of field trials. The main conclusions of the project were:

- 1) Crop simulation modelling technology has matured to the point where it can add value to plant breeding programs targeting dryland production systems in a number of ways including
  - a) Helping breeders to understand the types of environments targeted by their breeding program.
  - b) Determining the long term frequency of the different types of environments in the target region.
  - c) Assigning an environment type to a particular breeding trial and using that type as a way of dealing with GxE interactions either by weighted selection or selecting for specific adaptation to particular types of environments
  - d) Identifying best bet combinations of traits and management systems for different environments.
  - e) Using modelling frameworks to dissect complex traits into simpler component traits that are more amenable to conventional or marker assisted selection.
  
- 2) It is feasible to detect marker trait associations for quantitative traits in breeding populations using whole marker scans. This approach can add value to breeding programs by
  - a) Increased relevance of the marker trait associations to target breeding populations.
  - b) Enhancing the capacity to identify multiple alleles and identify the best alleles for selection.
  - c) Increased efficiency of implementation of marker assisted selection via reduced need for validation experiments, reduced genotyping of lines not associated with material that is likely to contribute to varieties and the capacity to use genotypic information for QTL detection and selection.
  - d) The potential to allow information from genome sequencing projects to be used directly in applied breeding programs.
  
- 3) Appropriate informatics systems that can effectively store and retrieve large quantities of data combined with analytical tools to allow this information to be used in selection decisions are essential to the implementation of both sets of technologies.
  
- 4) There are substantial opportunities for synergistic benefits to flow from combining crop simulation modelling with whole genome QTL analysis in breeding populations.

## Recommendations

Research in this project indicates that both crop simulation modelling and QTL detection in plant breeding programs have significant potential to improve the rate of genetic gain in plant breeding programs. As such they warrant additional research investment.

### QTL detection in breeding populations

1. Further develop or adapt statistical methods for detecting QTL in breeding populations.
2. Develop improved informatics systems such that data from a whole genome marker screening of breeding populations can be used for trial analysis in a cost timely manner.

### Simulation

1. Extend the simulation capacity now available in sorghum to other important crops where drought adaptation is a major target of selection.
  2. Enhance our capacity to carry out detailed trait dissection using gene to phenotype modelling.
-

3. Develop improved ways of characterising trial environments and converting them to a form that can be used in simulation such that modelling can be used as part of a plant breeder's suite of tools. Current characterisation techniques are labour intensive and therefore expensive. However it should be relatively simple to identify improved methods of accessing the same information. This might include the use of automated soil moisture measurement systems and biomass and leaf area measurement via remote sensing. Remote sensing offers opportunities to reduce the requirements for both sampling tasks. It also provides opportunities to inexpensively collect a wealth of additional data, which can be used to increase the rate of genetic gain achieved from the current investment in field testing. Currently our capacity to take additional measurements in field trials is limited by the costs associated with additional data capture, principally the need for human involvement. This has led to only a few measurements being taken at critical times (e.g. grain yield, maturity and disease ratings). These measurements capture the integrated response of individual entries to the complex interaction of environmental factors (soil type, rainfall temperature and humidity etc), management factors (planting date, fertility, population row spacing etc), spatial variability, and genetic variation in plant responses. The integrated nature of these measures means that genetic variation in crop development characteristics, such as canopy leaf area development, is often poorly captured and not exploited. As a major controller of genetic variation in water use patterns, variation in leaf area development is a critical component of enhancing genetic progress for drought adaptation. If this data could be captured cheaply via remote sensing and integrated via simulation it has the potential to greatly enhance progress in breeding programs.

## Outcomes

This project proposed two major outcomes

1. Development of a novel integrated crop improvement system with potential application to a range of GRDC funded crop improvement programs. In the medium term this would lead to greater yield advance in these programs and deliver superior varieties to growers.
2. Enhancement in the rate of improvement in sorghum yield, particularly the development of germplasm adapted to wide row management systems.

### Economic outcomes

Improved practices and genetic improvement are key drivers for improving productivity and profitability of grain growers. The expected outcome of this project is an increase in the rate of genetic gain in crop improvement programs focused on complex traits such as grain yield and drought adaptation. This was to be achieved by demonstrating the extra value associated with integrating a range of new technologies into an active crop improvement program and exploiting their synergistic interactions.

Plant breeding has yet to take full advantage of the dramatic technological changes that have occurred in recent years. To a large extent, genetic gain has been maintained by increasing the size and/or number of breeding trials. This project aimed to increase the information generated by each trial by implementing two new technologies, which are declining rapidly in cost (cf field trials). These technologies also have the potential to generate synergistic benefits. The project examined the use of 1) whole genome marker screening to detect QTL in breeding populations, and 2) simulation modelling to enhance the understanding of complex traits and GxE interactions in breeding trials.

For 1) the project showed that it was feasible to detect regions of the genome affecting quantitative traits in breeding populations. This approach increases the relevance of marker trait associations to breeding populations, allows multiple alleles to be detected, and speeds the time to marker implementation by removing the validation step. This type of approach could change the way breeding for quantitative traits is carried out in future.

For 2) the project used simulation modelling to explain significant amounts of the GxE observed in breeding trials. This allowed better identification of the types of environments experienced by trials and their frequency in the environments targeted by the breeding program. Simulation was also used to identify combinations of traits that suit particular types of environment and management systems using the example of skip row sorghum. We now have a breeding program with a prototypical design for the integration of simulation modelling.

The combination of 1) and 2) was also investigated as was the informatics support required to integrate the information. We consider that the project has demonstrated the opportunities for considerable improvement in the rate of gain in crop improvement programs by integrating new technologies.

## Achievement/Benefit

The objective of the project was to explore ways of helping crop improvement programs use new technologies to improve progress for complex traits. The technologies included marker assisted breeding, simulation modelling, increased computer processing speed, informatics and enhanced statistical analysis. Despite their great potential, we believe that these new technologies are yet to deliver the large advances in efficiency in plant breeding programs focused on quantitative traits that they promise. We contend that the best returns to crop improvement programs, and therefore to grain growers will occur when they are integrated into a redesigned crop improvement system rather than being applied in an uncoordinated way to an existing plant breeding program.

One of the main drivers for this research was the idea of adding value to increasingly expensive field based phenotyping for yield using technologies that are either low cost or are declining in cost. To this end the main areas of research were 1) developing the capacity to detect QTL in breeding populations and 2) using simulation modelling to enhance plant breeder's understanding of complex traits and GxE. Both of these technologies are becoming less expensive and can be used to extract additional value from field trials.

The research described in this report makes a significant contribution to the thinking around integrating these technologies into breeding programs by attempting to deploy these technologies in an active plant breeding program. The research highlighted practical constraints particularly in informatics and analysis that require additional research investment. The project produced a large number of outcomes across a diverse range of discipline areas either directly or by collaboration. The results of the project demonstrate the impact that effective integration across disciplines can achieve.

### 1) Development of structural genomic and mapping resources in sorghum and other species

The project developed and adapted improved methods of genetic map construction and validated the effectiveness of these methods initially using simulated data sets, and then using data from mapped and sequenced mapped markers. This work highlighted significant problems with the capacity of software widely used in Australia for map construction to generate accurate marker orders. Incorrect marker ordering has a significant impact on the effectiveness of QTL detection and marker assisted selection. These findings were shared with other GRDC funded crop researchers in a variety of formal and informal communications.

This project contributed to the development of the DArT technology for sorghum by DArT Pty Ltd via the provision of DNA from diverse germplasm, producing genetic maps using DArT and by providing beta testing for the technology for a variety of applications.

The project developed one of the world's best genetic maps of sorghum (2100 markers) by integrating data from six genetic maps. This was achieved using local mapping populations and by developing collaborations with scientists in the USA (TAMU) and France (CIRAD). The map construction activity was led from within this project.

The map produced is now linked to the sorghum genome sequence by access to the sequence of DArT markers (via the collaboration with CIRAD and DArT) combined with sequence mapping of 213 SSR markers and 42 RFLP markers on the map.

This integration of the genetic map and genome sequence combined with access to cost effective whole genome scans (eg DArT) will lead to a revolution in the way markers and genome information is used in breeding and pre-breeding applications. This combination of technologies makes previously interesting data from a variety of sources available for use in applied breeding in a cost effective and timely manner. For example it is now a trivial issue to take information on candidate genes for traits identified in another species and identify their location in the sorghum genome and relate that to any QTL information available in sorghum. Similarly, all of the genes located near a QTL can be identified and assessed as candidate genes for the trait. At a more applied level the consensus map and genome sequence can be used to rapidly and cheaply develop polymerase chain reaction (PCR) based markers (e.g. simple sequence repeat (SSR)) in regions of the genome containing QTL for key traits. It can also be used to quickly identify

the location of QTLs identified in new or historical QTL mapping studies in sorghum (and other grass species) in breeding populations.

## 2) Demonstration of the feasibility of detecting QTL in breeding populations

The detection of marker trait associations has typically been done in populations specifically developed for this purpose (recombinant inbred lines (RIL), doubled haploid (DH) etc). This approach has proved relatively effective for traits under relatively simple genetic control such as disease resistance. The approach has a number of disadvantages if the objective is to identify QTLs for marker assisted selection. For example, this approach invests large amounts of resources (genotyping and phenotyping) into lines which are unlikely to become varieties. Markers developed in experimental populations may not be effective in breeding populations and a second phase of marker validation is required before marker assisted selection is commenced. The multiphase approach is time consuming and it is often many years before markers can be used in breeding.

This project demonstrated the feasibility of detecting marker trait associations in breeding populations, which had previously been selected for agronomic type. In the case of the proof of concept example of flowering time, the approach used detected many more QTLs than would typically be found in a single mapping population. Comparison to published literature and candidate genes indicated that the majority of these marker trait associations were real.

This approach to breeding is still in its embryonic stage. The QTL detection methods require considerable refinement as does our capacity to handle the large quantities of data effectively. The availability of this data makes possible a number of promising changes to the way plant breeding programs operate.

The sorghum breeding program is already beginning to use the data from this research in parental selection and aims to start a program of marker assisted recurrent selection to pyramid favourable regions affecting yield. In the future we believe that rather than using the marker trait associations for selection in other populations they can be used directly to improve the confidence we have in the superiority of lines in trials in much the same way as pedigree data is being used in animal breeding.

## 3) Investigation of the role of simulation modelling for use in environment characterisation, GxE analysis and trait dissection

Genotype x environment x management (GxExM) interactions are a major limitation to genetic progress in cereal breeding programs targeting dryland cropping systems. In particular variation in the timing and severity of water stress leads to re-ranking of varieties in trials from location to location. This re-ranking makes it difficult for breeders to identify superior varieties. GxExM interactions are the major factor driving the need for expensive multi-environment trials which are the major cost of most plant breeding programs.

This project demonstrated that simulation-based environment characterisation has the potential to explain GxE interactions caused by differences in the timing and intensity of water stress associated with different breeding trials and for entries within trials. Using this technology it was possible to place environments sampled in the context of their relevance to the range of environments targeted by the program. This capability was applied in a manner that could be used as an aid to selection decisions.

The modelling capacity also facilitated dissection of complex traits, such as yield, into less complex sub-traits that are more amenable to selection and identification of associated genomic regions. The emerging capacity in gene-to-phenotype modelling and prediction enabled a putative evaluation of particular sets of traits in the environments targeted by the breeding program before selection is commenced.

## 4) Identification of combinations of sorghum traits suitable for skip row sorghum

The recent change to wide or skip row sorghum production systems has occurred with little or no basic research into the types of genotype characteristics required to achieve best results from these systems. Agronomic research to date has consisted of choosing the best of available varieties. It is likely the skip row systems will produce optimal results only when they are used in conjunction with hybrids specifically bred for these systems. Due to resource constraints, it is difficult for plant breeding programs to run breeding trials aimed at both skip and solid planting arrangements. Part of this project involved using simulation and field trial analysis to identify combinations of traits that are likely to achieve superior performance in skip row systems. Comprehensive simulation studies across the sorghum

production region identified the possibility for genetic regulation of tillering as a means to improve production risk. The greater control of canopy development afforded by this genetic manipulation, allowed useful adaptation of density and row configuration patterns, especially in marginal environments. This information has been transferred to public and private sector sorghum breeders and has resulted in changes to selection programs (e.g. the development of low tillering sorghum).

5) Contributed to attracting significant investment from private sector seed companies into sorghum trait analysis and modelling.

GRDC's investment in this project contributed to securing approx \$1.8M over four years for an Australian Research Council (ARC) linkage project with partners the University of Queensland (UQ), Department of Primary Industries and Fisheries (DPI&F) and Pioneer Hi-Bred on "Modelling gene-to-phenotype relationships using crop growth models". The linkage project involved complex adaptive trait analysis and modelling in maize and sorghum. By providing access to populations, marker data, materials and molecular genetics expertise developed in the integrated breeding technologies project, mutually beneficial synergies were realised. Successful outcomes of the ARC project of major relevance to the Australian sorghum industry included enhanced physiological and genetic understanding of sorghum traits including grain size, root architecture, stay-green and partitioning. In addition, enhancements to the crop modelling capability for sorghum were of direct benefit to aspects of the integrated technologies research.

6) Communications outcomes

Project staff communicated results and ideas generated throughout the course of the project to industry and researchers. In many cases (e.g. improved mapping techniques, traits for skip row sorghum) these communications led directly to improved outcomes from grains research. The final workshop was well received by those who attended and strong feedback ratings suggested that the ideas presented will influence the research approaches taken by the scientists who attended.

## Other Research

Following presentations in the US and Europe based on this project and the associated ARC project, Professor Hammer has been invited to participate as a work program leader in an European Union (EU) Framework (FP7) large grant project application being coordinated by Dr Francois Tardieu (INRA, Montpellier, France). The project addresses the topic of water stress tolerance and water use efficiency in food crops. The proposal is focused on physiology and genetics of stress adaptive traits in wheat and maize (and possibly sorghum). The focal traits are grain abortion, vegetative growth maintenance, root system efficiency, and transpiration efficiency. Prof Hammer will be leading some aspects of the gene-to-phenotype modelling program and sorghum component. EU large grants are funded to a maximum of €6M over 4-5 years. While none of these funds can be allocated to participating Australian scientists, there is possibility for funding of linked Australian research via the International Science Linkages program of the Australian Government. The generic cereal crop model in APSIM, which has been developed based on sorghum research, will be utilised and significantly enhanced via this project (should it be successful). There is considerable opportunity to enhance the role of modelling in crop improvement by development of this technology.

Following the development of the sorghum consensus map, anchored to the physical sequence map, new opportunities have arisen to identify candidate genes for key target traits, either through identifying homologous genes in sorghum from candidate genes in other species, or through a positional cloning approach by identifying predicted genes in fine-mapped QTL regions. Current prioritised traits of interest for candidate gene identification include stay-green, midge resistance and fertility restoration. Further work to validate the candidate genes identified could involve an Eco-TILLING reverse genetics strategy, to look for naturally occurring variation in selected candidate genes and to then relate the observed genotype variation to the phenotype. With the inclusion of the diversified gene pool (DGP) parental lines in the eco-TILLING set, this would allow rapid validation of candidate genes for QTLs identified across the DGP populations. The high degree of genomic synteny between cereal crop species in the Poaceae family could also be utilised in a comparative phenomics and genomics strategy to identify genes associated with high yield in water limited environments and also to transfer this knowledge to barley and wheat. Comparative phenomics offers opportunities to use the advances made in physiological understanding of key traits in one crop, in the application to other crops. It offers additional opportunities for insights into different crop responses to the same environmental trigger, such as drought, which have evolved over millennia. We envisage that comparative phenomics will offer oppor-

tunities to not only rapidly identify many potentially valuable traits to pursue, but also to advance the genetic improvement of key selected traits, through the application of the physiological dissection of quantitative traits in one species across other related cereal crops, combined with the exploitation of the range of genetic material and inherent adaptation present within cereal crops.

## Intellectual Property Summary

The outputs of this project will generate most benefit to growers if they are made freely available. No attempt is being made to protect any of the project outputs with the exception of germplasm developed as part of the DPI&F sorghum core breeding program. This germplasm will be licensed under existing commercialisation arrangements.

## Collaboration Organisations

Via association with related projects (CGIAR-GCP and ARC) activities in this project have been linked internationally with -

Dr M. Cooper, Dr C Messina, Dr Z Dong, Pioneer Hi-Bred International, Johnston, Iowa USA  
Dr F Tardieu, Dr C Welcker, INRA, Montpellier, France

Through the consensus map research activities, this project has been internationally linked with:

Jean-Francois Rami, Sophie Bouchet based at CIRAD UMR DAP, TA A-96/03, Av Agropolis, 34398 Montpellier CE-DEX 5, FRANCE

Patricia E Klein based at Department of Horticulture and Institute for Plant Genomics and Biotechnology, Texas A&M University, College Station, TX 77843-2123, USA

Robert R Klein based at USDA-ARS, Southern Plains Agricultural Research Center, College Station, TX 77845, USA

## Additional Information

Further work is being carried out in Project DAQ00132.

---