



Australian Government

Cotton Research and
Development Corporation

FINAL REPORT 2008

*If you are participating in the presentations this year, please provide
a written report and a copy of your final report presentation
by 31 October.*

If not, please provide a written report by 30 September.

Part 1 - Summary Details

Please use your TAB key to complete Parts 1 & 2.

CRDC Project Number: 05CSP009

**Project Title: Cotton biotechnology: innovative genetic
solutions to enhance cotton productivity**

Project Commencement Date: 1 July, 2007 **Project Completion Date:** 30 June, 2008

CRDC Program: 5 Breeding and Biotechnology

Part 2 – Contact Details

Administrator: Tracey Williams
Organisation: CSIRO Plant Industry
Postal Address: P.O. Box 1600 Canberra ACT 2601
Ph: 02-62465281 **Fax:** 02-62465000 **E-mail:** Tracey.Williams@csiro.au

Principal Researcher: Danny LLEWELLYN
Organisation: CSIRO Plant Industry
Postal Address: P.O. Box 1600 Canberra ACT 2601
Ph: 02-62465470 **Fax:** 02-62465000 **E-mail:** Danny.llewellyn@csiro.au

Supervisor: Danny Llewellyn
Organisation: As above
Postal Address:
Ph: **Fax:** **E-mail:**

Signature of Research Provider Representative: _____

Part 3 – Final Report Guide (due 31 October 2007)

(The points below are to be used as a guideline when completing your final report.)

Background

1. Outline the background to the project.

CSIRO has established an effective breeding program that continues to deliver high yielding and high performing varieties containing transgenic traits for more sustainable and profitable insect and weed control. To stay competitive globally, new varieties must be developed more rapidly and efficiently and this can only be achieved through the adoption of biotechnology and modern molecular marker techniques. Over the last 15 years the Canberra Biotech team has augmented the efforts of CSIRO's breeders to allow the introduction of Monsanto's Ingard, Roundup Ready, and later Bollgard II and Roundup Ready Flex traits into elite germplasm that has had a significant impact on the economics and environmental footprint of cotton. As the conventional variety suite is expanded each new variety must be converted by repeated backcrossing to BG II/RRF (or any new transgenic trait introduced into the program) and this requires extensive molecular screening to follow the traits (all three genes segregate independently) through segregating populations. This is carried out in Canberra by high throughput biochemical and DNA screening methods that detect the Bt protein or site of insertion of the transgene, effectively molecular markers for each of the traits. Molecular markers for disease, yield or quality determinants could be used in a similar way in our conventional breeding to reduce population sizes before extensive field testing is needed and would accelerate CSIRO's output if the right markers could be found. Some markers reportedly linked to good fibre quality have been published, but we need to establish that they are relevant to our germplasm. Yield is also critical to cotton and if we had a better understanding of the genetic determinants of yield we might more effectively select for varieties with both high yield and quality or manipulate yield by GM to keep Australian cotton ahead of its competitors.

Objectives

2. List the project objectives and the extent to which these have been achieved.

The project aims to augment the capacity of CSIRO's breeding team to produce high yielding high quality transgenic varieties with innovative and competitive performance through the use of molecular screening tools that will allow them to follow transgenic or other genetic traits through breeding populations at a greater speed and with greater reliability than through conventional selection.

The project originally had three broad objectives: the identification of novel traits or markers for enhanced fibre yield through a better understanding and manipulating of the genetic controls of seed number in cotton (20% effort); the use of molecular screening methods to support the incorporation by CSIRO cotton breeders of Monsanto Bollgard II and RRflex traits (and any new licensed traits for insect and herbicide resistance (eg. BG III) or for improved agronomic performance that becomes part of the breeding effort (eg. drought tolerance)) into new elite conventional CSIRO germplasm (70% effort); and preliminary investigations into the utility of published molecular markers linked to fibre traits such as length, strength and fineness to enhance the speed of development of elite conventional and transgenic Australian varieties with both high yield and quality (10% effort).

Because of changes in the funding relationships between CSIRO and CRDC the original three year project had to be reduced in scale to one year with a subsequent reduction in milestones and potential for achievement. The main objective remained to continue working closely with the breeding program to screen early generation transgenic breeding lines and this was successfully achieved with contributions being made to the 2007-2008 Single Plant Selection screening of both Bollgard, Roundup Ready Flex and BG/RRflex stacked material (for presence and absence of genes and their zygosity) as well as screening of the next generation of the previous years SPS selections to ensure that they were indeed homozygous. Over 300,000 ELISAs, 15,000 DNA preps and 40,000 PCR reactions were carried out by the technical team and these were essential for the advancement of that material into subsequent evaluation trials. This took most of the time of the technical team but they

were also able to make some small progress in the other two goals of evaluating the genetics of seed number in cotton and investigating published molecular markers for fibre quality traits. A genetic population of a cross between the low and high seed number species of cotton (*G. barbadense* and *G. hirsutum*, respectively) was produced and over 200 plants phenotyped in the F2 generation for ovule number per boll. This established that seed number was heritable and probably conferred by a small number of genes (perhaps as little as 1 or 2 major genes). In the molecular marker area a review of the literature was carried out and markers examined from different groups. Unfortunately the different groups used different marker systems, markers and populations so it was not possible to compare between them directly as they did not use any common markers. Time was invested in setting up the experimental gel systems to look at microsatellite markers on polyacrylamide gels and this is now working reasonably well. A small number of markers reported linked to fibre strength and length were examined in different Australian cotton genotypes but as all of them were developed from inter-specific crosses none were polymorphic between different Australian *G. hirsutum* elite lines tested. Considerably more investments in time would be needed to properly evaluate the markers and to screen Australian material for DNA polymorphisms in the same genomic regions that might be useful as markers for breeding. Since the initial investigation was done more effort is being made by the International Cotton Genome Initiative to align DNA marker maps developed by different groups. The potential to explore the use of markers for fibre quality remains to be investigated in future projects.

Methods

3. Detail the methodology and justify the methodology used. Include any discoveries in methods that may benefit other related research.

CSIRO has expertise in the use of molecular markers in transgenic breeding using both biochemical and DNA markers to track Monsanto insecticidal and herbicide tolerance transgenes through breeding populations of up to 100,000 plants each year. Our breeders backcross every new conventional variety with an existing BG II/RRF variety, but in order to recover the unique genetic makeup of the original conventional cultivar large populations (several thousand per cross) need to be screened using high throughput ELISA and DNA screens in Canberra to identify sufficient material from which to make their selections in later generations based on agronomic performance. ELISA assays are used to detect the Bt proteins in extracts of leaves or seeds and provide rapid and sensitive detection of transgene containing plants (RRF plants are selected in the glasshouse or field based on tolerance to Roundup herbicide), but only sensitive PCR methods can be used to detect each of the three transgenes AND their zygosity (whether they will breed true in later generations).

PCR screening quickly reduces population sizes down to only those plants that are homozygous for all three genes and saves the breeders both time and land in later selections. PCR tests used in BG II/RRF breeding are DNA markers for the GM traits and their effective use demonstrates that markers for other traits could also be usefully applied in the breeding program. If the appropriate markers were available it could accelerate the rate of advance for more complex genetic traits like yield, quality or disease tolerance. Progress here is currently constrained by either difficult screening requirements for traits like fusarium wilt tolerance or the complex, often negative interactions between different yield and quality parameters. Breeders must balance workable population sizes with the consequent genetic gains across a variety of yield and quality parameters. Markers for fibre properties like length, strength and fineness have been identified in Chinese, Brazilian and US varieties, but their reliability is unclear (few coincide between different varieties) and they may not be applicable to Australian germplasm.

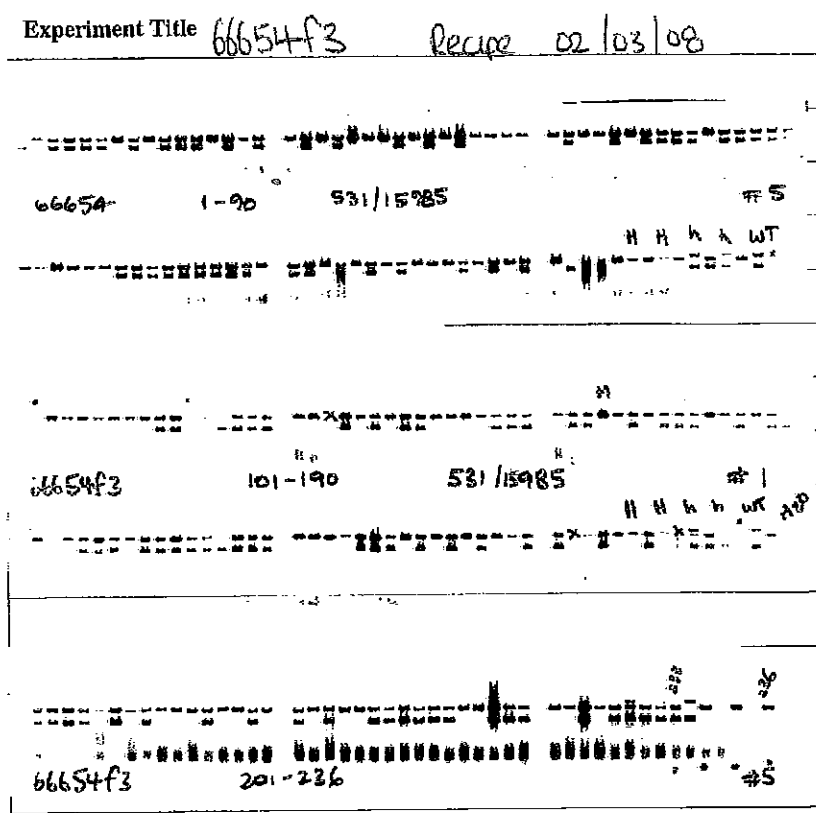
Results

4. Detail and discuss the results for each objective including the statistical analysis of results.

Transgenic Screening Activities:

2007-2008 Field Screening

The major activity each year was screening plants from field plots of Single Plant Selection (SPS) populations of elite breeding lines backcrossed to the various Monsanto transgenes for insect (MON531 and MON15985) and herbicide (MON88913) tolerance. In 2007 13 Bollgard II, 1 RRflex and 17 Bollgard II/RRflex stacks and 2 Liberty link/Bollgard II and 2 Bollgard II Pima families were screened. Each Family involved screening F2 segregating progeny of backcrosses between lines carrying the transgene(s) and new conventional cultivars being converted to a transgenic version. Each family involved screening between 500 to 6,000 individual F2 plants (higher numbers when there were more segregating genes). A preliminary screen was carried out on cotyledon samples using ELISA assays to remove plants lacking both insecticidal genes and then the remaining plants were sampled and leaves freeze dried to be extracted for DNA for PCR testing to determine the zygosity of the different transgenes. Where possible PCR tests were combined to reduce sample processing and remove potential OHS issues for the technical staff carrying out many hundreds of thousands of repetitive operations each season. A typical screening gel to determine the zygosity of the two Bt genes in the 66654 family is shown below.



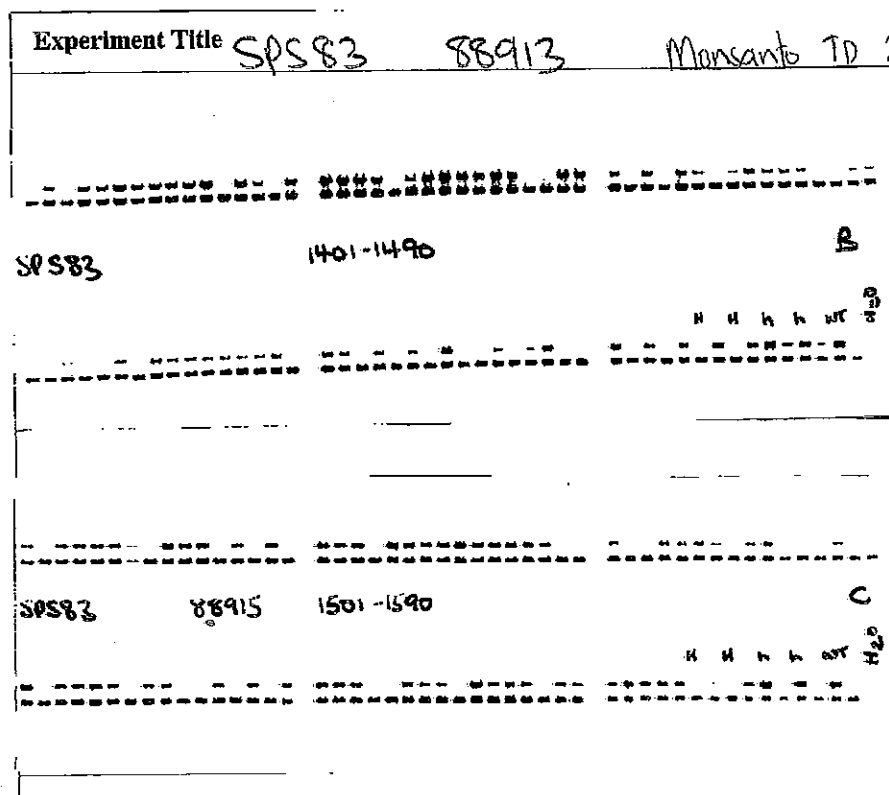
Plant homozygous for both insecticidal genes (Bollgard II) contain only a single upper band in this multiplexed assay.

Many hundreds of these gels were run during the season to identify homozygous plants that were then left in the field to be harvest during April 2008. Similar gels were run to determine the zygosity of the RRflex gene.

2007-08 Glasshouse Screening activities:

Throughout the year the breeding team were also developing new populations from crosses between newly selected conventional varieties and transgenic varieties already containing the Monsanto or Bayer traits. At each generation in backcrossing the plants were screened by ELISAs to select for plants containing the two Bollgard II genes and selected plants taken on for a further backcross or selfed to generate the F2 families. 12 Bollgard II families, 14 BollgardII/RRF and 1 RRF were generated and screened. In order to save time a number of

these were taken through to F2 lines in the glasshouse rather than the field and had to be screened for zygosity of the various genes. This material will go directly into the field in 2008 as progeny rows saving both time and field resources.



Typical Screen for zygosity of the MON88913 RRFlex gene in the SPS83 family.
Homozygous lines have only a single lower band

Winter Screening Activities 2007-2008:

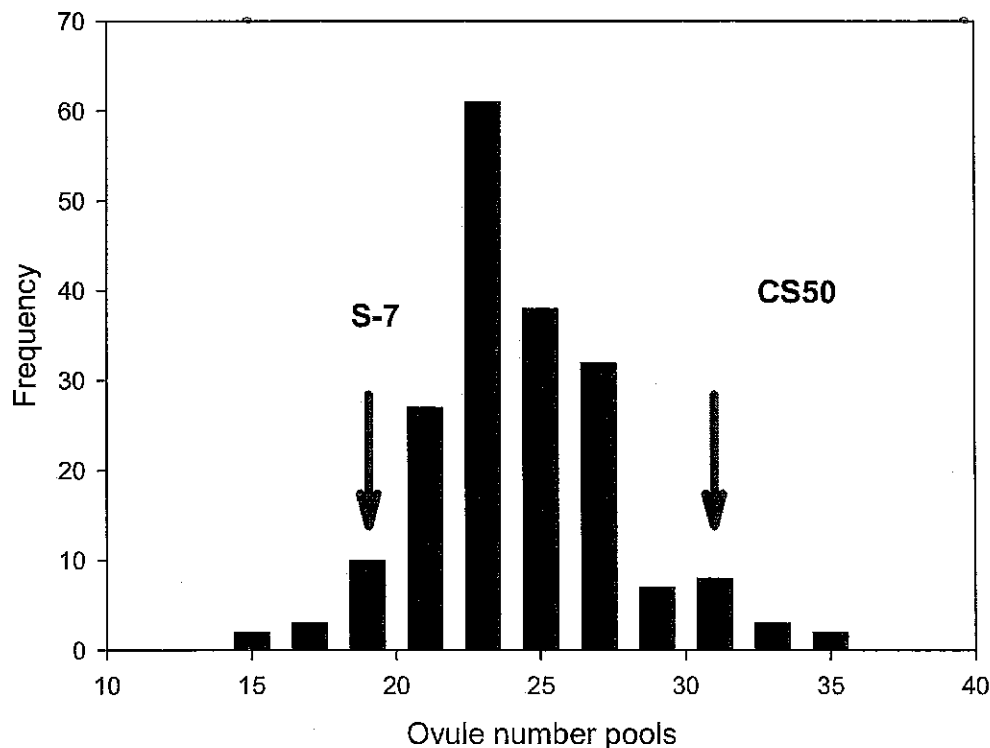
During the winter months each year, after the previous 2006-07 season's SPS families had been harvested, ginned and delinted in Narrabri, there was still screening work that had to be carried out. This was primarily a confirmation of the zygosity testing of SPS lines identified in the previous year by testing the seed harvested from those plants. This involved carrying out ELISA assays for the two Monsanto insecticidal traits in Bollgard II lines on seed extracts from 24 seeds of each plant identified as being homozygous. The DNA testing is about 95% reliable so there are always some lines that are either incorrectly scored or incorrectly harvested due to human error. This additional screen removes these lines and ensures that the progeny rows planted out in the following season for agronomic evaluation are true breeding (since between 150-400 homozygous lines are advanced in each family this represents a considerable amount of seed testing that was carried out – a few hundred thousand individual ELISA assays). Similar activities were also required for the 2007-08 SPS lines over the winter of 2008 prior to the next planting season and these are on going. This testing prevents segregating lines being planted out with true breeding lines and having the true breeding lines being cross contaminated with pollen not containing all of the appropriate genes. During the summer and winter there was also a smaller amount of screening from glasshouse plants being grown for new crosses set up by the breeders (involving several thousand ELISA and DNA screens each few months).

Large scale screening activities were therefore required relatively continuously throughout the year from the technical staff with smaller gaps of less activity just prior to planting time each year (except for the glasshouse screening of new breeding families that can occur at any time of the year). This allowed only a small amount of time to explore the two other research activities nominated in the initial project proposal.

Genetics of Cotton Seed Number.

Genetic crosses had been set up between a *G. hirsutum* line CS50 and a *G. barbadense* line Pima S-7 that differed in seed number per boll based on commercial seed data. These lines were chosen as we had some prior information on DNA marker polymorphism between the species from an earlier PhD project on molecular markers with Dainis Rungis. Data from several plants and averaged across five flowers per plant indicated that the *hirsutum* line had on average 30.6 ovules per flower in four-five locules while the *barbadense* line has 18.9 ovules per flower in three locules. The F1 progeny had on average 22.3 ovules per flower suggesting that there may be semi-dominant repressors of high ovule number in cotton. Over the autumn of 2008 approximately 200 F2 plants were grown up and five flowers from each plant dissected and ovule numbers determined (ovule numbers rather than seeds produced were counted to exclude differences in fertilisation efficiency that can be complicated in inter-specific crosses).

F2 Segregation - Ovule Number
CS50 x Pima S7



Histogram of the segregation of ovule number across the F2 population. Arrows indicate the ovule number of the parental genotypes. Transgressive segregation was observed at each end of the distribution.

The frequency distribution of ovule number followed an approximately normal distribution, suggesting that there were not a large number of genes conferring ovule number. Calculations of the number of genes conferring a trait using the Wright-Castle index (using the variance in ovule number of the F1) gives a value of 1.8 genes (or 1.4 genes using the average variance in ovule number of the two parents, an alternative measure of the number of genes). This indicates that ovule number might be conferred by 1-2 genes, so might be a good target for QTL mapping and identification of a specific gene(s). DNA has been collected from all the plants and seed maintained from plants at the two extremes of the distribution to see if high or low ovule numbers are inherited. Recent studies in rice using such an approach identified a

single major QTL conferring panicle number that because of the availability of a genome sequence could be narrowed down to differences in expression of a cytokinin oxidase gene that could be manipulated using transgenic plants to increase panicle number. While the two systems are not equivalent our data suggests that a similar approach could work with cotton but would require a much bigger investment in time and resources. Progress would be hampered until there was a good genome sequence and integrated genetic maps for cotton but this is likely to be available over the next few years. Again as the project was only for one year we could not take it any further.

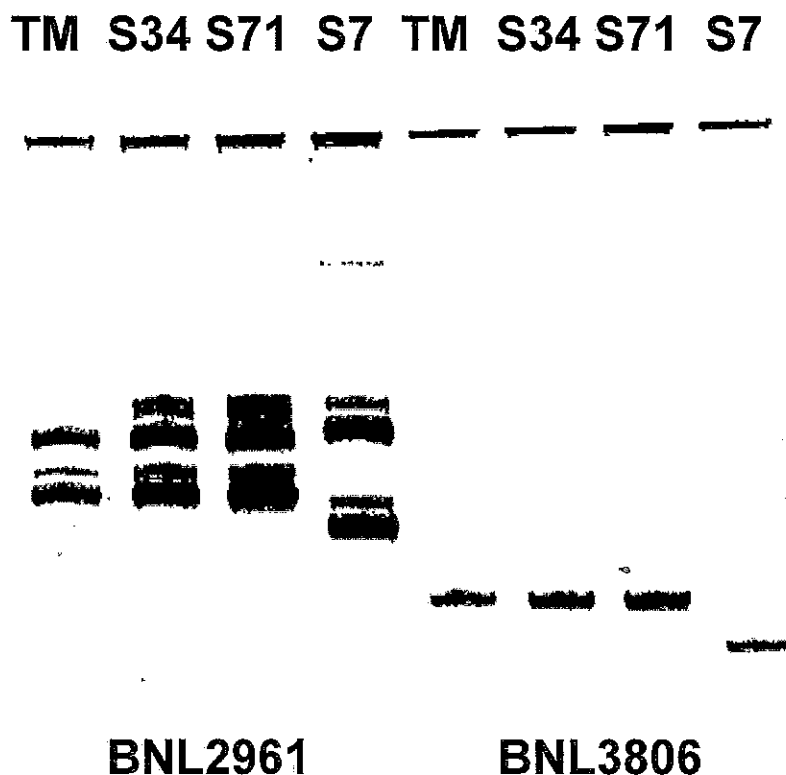
Cotton Fibre Quality Markers

Over the past few years there has been renewed effort in applying molecular markers in cotton to map QTLs for important agronomic traits like fibre quality. The major groups are CIRAD in France, Huazhong Agricultural University and Nanjing University in China, University of Georgia, Texas A&M University and USDA-ARS in the USA all of whom have published genetic maps from *G. hirsutum* x *G. barbadense* crosses (inter-specific crosses are needed in cotton to achieve sufficiently high levels of DNA polymorphism for mapping) and identified a large number of QTLs for the various fibre quality traits measured by HVI. The literature was reviewed but it was soon apparent that there was:

1. a very large number of QTLs for each fibre trait like length, strength or fineness (a dozen or more for each quality trait in each different mapping population) and each conferred only a small percentage of the total variation for those traits
2. very little concordance between the different maps (compounded by different nomenclatures for the different cotton chromosomes)
3. very little ability to cross compare markers because each group has used their own set of markers and have not run any common markers that would allow the maps to be integrated.

Progress is now starting to be made with these different studies and co-ordination is just starting to be achieved through the International Cotton Genome Initiative. Different labs are being encouraged to run common markers (often the BNL microsatellite markers that were the first to be developed for cotton or the RFLP markers developed by Paterson to make the first fully resolved genetic map of cotton and now being converted to simple PCR markers). Some consensus in numbering the cotton chromosomes has also been achieved with the first 13 being from the A genome and the second 13 from the D-genome and homeologous chromosome pairs being given appropriate numbers but many of the earlier studies will have to be reassessed to conform with this numbering. We began to set up high throughput acrylamide gel systems to examine various microsatellite markers (PCR markers to detect small differences in the numbers of repeated sequences within specified genomic regions) and now have this going reasonably well. A couple of the BNL markers were selected from around the regions identified by some of the Chinese groups and although these gave size differences between pima and upland varieties they were invariably the same size in a number of different Australian breeding lines tested. Considerable more time and effort would need to be expended to find polymorphisms in these regions that could be used to screen for segregation with fibre traits but we believe that we have developed the right capabilities in this area to be able to do that if we were confident a particular genomic region conferred a reasonable proportion of the genetic variation for particular fibre traits of importance. Some of the BNL, CIR and NAU markers (used by the US, French and Chinese labs, respectively) are from expressed genes so it would be possible to sequence the same ESTs from different Australian cultivars to be used in crosses and develop new markers from that or to screen BAC libraries with the microsatellite probes and sequence around the marker to obtain a big enough region that could be sequenced from different Australian varieties to develop a new marker in the same region. This is obviously very labour intensive and would

only be warranted for a marker linked to a genomic region conferring a large contribution to a particular fibre trait.



DNA marker profiles for two markers linked to fibre strength (BNL2961 on Chr 10) and Length (BNL3806 on Chr 25). TM, *G. hirsutum* TM-1, S34, Sicala 34; S71 Sicot 71 and S7 *G. barbadense* Pima S-7. Note BNL3806 markers are all identical in the upland varieties but a different size or pattern in the pima variety.

Outcomes

5. Describe how the project's outputs will contribute to the planned outcomes identified in the project application. Describe the planned outcomes achieved to date.

The major planned outputs were screening data that contributed towards the early generation transgenic variety development by the CSIRO breeding program. This was achieved and about a dozen SPS populations were allowed to progress to the next stage of selection by the breeding team. Some of these plants and lines will eventually end up as new Bollgard II, RR Flex and stacked cultivars released to Cotton Seed Distributors in about four years time and to the Industry in about six or seven years time. The second output was genetic data on a potential new trait for increasing seed numbers and hence potentially fibre yield in Pima or Upland varieties. The genetic nature of seed number was shown to be reasonable simple and heritable so this could open up more research into understanding the molecular basis for seed number and the possibility of increasing seed number through breeding or transgenic approaches. Finally we had hoped to explore the possibility of using markers for fibre quality in cotton breeding but the low level of DNA polymorphism in cotton and the inability to compare different genetic maps means that this will have to be further down the track once the cotton community puts more effort into combining genetic maps from different laboratories (using

common sets of markers). We believe that this will happen eventually and will be encouraged by the web based resources in cotton markers being developed under the auspices of the ICGI (eg Cotton Marker Database: <http://www.cottonmarker.org/>).

6. Please describe any:-
- a) technical advances achieved (eg commercially significant developments, patents applied for or granted licenses, etc.); NONE
 - b) other information developed from research (eg discoveries in methodology, equipment design, etc.); NONE
- and
- c) required changes to the Intellectual Property register. NONE

Conclusion

7. Provide an assessment of the likely impact of the results and conclusions of the research project for the cotton industry. What are the take home messages?

The project continued the work with early generation transgenic cotton breeding being carried out by the CSIRO core breeding team and allowed them to set up new populations and progress previously established populations along the breeding pipeline that will eventually result in the release of new transgenic varieties to the industry. Those varieties could not be produced without the molecular screening carried out in this project and the research and transgenic varieties will invariably result in higher yields, lower chemical usage and higher profits for the growers who will use them in about eight years time. The other key messages from the project were that seed number is simply genetically controlled and might be a useful target for breeding or biotechnology, but would require a dedicated project and considerable further investment to establish its utility in breeding either new Pima or Upland varieties. Molecular markers for fibre quality are still at a very early stage but should become more useful for breeding once the genetic maps developed by different research groups are combined and this should happen over the next couple of years through the International Cotton Genome Initiative (ICGI) of which CSIRO is a member.

Extension Opportunities

8. Detail a plan for the activities or other steps that may be taken:
- (a) to further develop or to exploit the project technology.
 - (b) for the future presentation and dissemination of the project outcomes.
 - (c) for future research.

Development of the breeding lines into cultivars will continue within the existing pipeline of the breeding program and will in due course be released to growers through CSIRO's exclusive licensee, CSD. CSD will be responsible for the extension of the varieties as part of their normal marketing processes. The other research was purely preliminary and will require further input before a proper project could be developed. This will occur within the Cotton Breeding Australia Joint Venture set up between CSIRO and CSD to carry out research that will contribute to the development of new CSIRO varieties or better methods of breeding new varieties. CSIRO hopes to host the 2010 Research meeting of the ICGI that will be a useful forum to progress both the sequencing of the cotton genome and integration of global mapping projects in cotton that will eventually allow us to quickly determine the underlying molecular determinants for many important agronomic traits.

9. A. List the publications arising from the research project and/or a publication plan.
(NB: Where possible, please provide a copy of any publication/s)

NONE

- B. Have you developed any online resources and what is the website address?

NONE

Part 4 – Final Report Executive Summary

This project had three main objectives: to continue making contributions to CSIRO's transgenic cotton breeding program through molecular screening of breeding lines for the presence and absence of specific traits from Monsanto and Bayer; a genetic analysis of the determination of seed number as one component of fibre yield in cotton; and a preliminary evaluation of the molecular marker literature for cotton fibre quality traits to explore the potential of this new technology in breeding for quality.

The major contribution of the project was to the advancement of early generation transgenic breeding material and several lines were screened and progressed in the breeding pipeline by one generation.

A second component of the project evaluated the genetic determination of seed or ovule number in cotton from a cross between a pima (low seed number) and Upland (high seed number) variety. The inheritance of seed number in the F2 generation suggested that seed number may be simply determined by a small number of genes and hence might be a suitable target for more detailed mapping or even cloning of the gene(s) responsible. This would however require a much bigger and long term investment and would be more likely to succeed once the full sequence of the cotton genome is known in the next couple of years.

The final minor aspect of the project was to examine the utility of DNA markers for fibre quality traits in Australian breeding material. This involved a survey of the literature and the preliminary screening of Australian material with any potential markers. Unfortunately although there are a number of recent publications on mapping of fibre quality traits like length, strength and fineness these proved to be quite complex and determined by several genes each with a relatively minor contribution to the trait. The use of in-house markers to generate the genetic maps by the different groups also meant that the genetic maps of the different groups could not be compared and they even had different naming conventions for the different cotton chromosomes. These issues are being addressed by the International Cotton Genome Initiative and should make comparing between groups simpler in the future. A few markers supposedly linked to fibre strength and length QTLs were tested on a couple of Australian cultivars but were found to be mostly uninformative because of the low level of DNA polymorphism between cotton varieties.

We therefore conclude that considerable more work would be required to properly evaluate such markers and we suspect that they may not be that useful for our breeding program as they only account for relatively small percentages of the determinants for those traits.