# CHARACTERISATION AND MAPPING OF THE *XCM* RESISTANCE LOCUS IN AUSTRALIAN COTTON CULTIVARS BY THE USE OF MOLECULAR MARKER TECHNIQUES

Dainis Rungis[1], Danny Llewellyn[1], Liz Dennis[1] and Bruce R. Lyon[2]
Australian Cotton Cooperative Research Centre
CSIRO Division of Plant Industry,
P.O. Box 1600, Canberra, ACT, 2601[1]
School of Biological Sciences A12,
The University of Sydney, Sydney, NSW, 2006[2]

## Introduction

Bacterial blight of cotton is caused by the bacterium *Xanthamonas campestris* pv. *malvacearum* (*Xcm*). This infection attacks leaves and bolls and causes early leaf drop and affects boll development and normal boll opening. In Australia the CSIRO breeding program has developed varieties that are resistant to *Xcm* infection. The use of these resistant cultivars has meant that the importance of this disease has been reduced. The resistance locus was introduced into the breeding program from the USA in the 1970's; however, the identity of the resistance gene or genes has not been fully clarified. *G. barbadense* or Pima varieties are susceptible to *Xcm*, and a breeding program is in place aimed at introgressing *Xcm* resistance into the Pima cottons from *G. hirsutum*. Further work is needed to characterise the source of the resistance, and possibly to develop molecular markers linked to the resistance locus in order to assist with breeding efforts. In order to map this resistance locus a mapping population was set up using backcrosses between the resistant *G. hirsutum* variety CS50 and the susceptible *G. barbadense* variety Pima S-7.

## Molecular markers

Markers can be used for a variety of applications. These include mapping of specific genes, cultivar identification, and studies of biodiversity. Initially, morphological markers and isozymes were used to perform these tasks, but these markers were not fully suited to these applications. There are not sufficient numbers of morphological markers and isozymes to provide detailed coverage of most genomes. In addition, these markers can be difficult to characterise, and do not show high levels of polymorphism or variation. Morphological markers may also be affected by environmental factors, and so are not entirely reliable.

Molecular markers are a more recent development, and possess many advantages which make them superior to morphological markers. Molecular markers can be used to directly examine the DNA in order to find polymorphisms in the sequence. This offers the possibility of finding large numbers of polymorphisms, or DNA sequence variants, which are not subject to environmental influences. There are several different types of molecular markers in use, and their suitability depends on the type of project being undertaken, and the facilities available. Each molecular marker technique has various advantages and disadvantages. These include the number of loci examined with each reaction, dominance or co-dominance and ease of application. Various marker

techniques detect different numbers of independent loci per reaction. The more loci detected per reaction, the fewer number of reactions that will be required to survey an equivalent amount of a genome. In addition, more than one allele may be detected at one locus. This is termed co-dominance. This means that a marker in a heterozygous state will be able to be distinguished from a homozygous locus. Again this improves the imformativeness of a particular marker. There are other factors such as the biological system to be studied, the availability of specific equipment and financial considerations that have to be taken into account when deciding between marker techniques.

A variety of molecular marker techniques were utilised within this project in order to search for markers linked to the *Xcm* resistance locus. These included AFLP's, SSR's and RFLP's. Each particular method has various advantages and disadvantages. The AFLP technique is useful as it can survey a large number of independent loci per reaction. However, these markers are dominant only, which means that at a given locus heterozygous individuals cannot be distinguished from homozygous individuals. SSR markers are co-dominant, but prior sequence knowledge of the genome is required, and so this technique can take a long time to set up, and must be tailored to each individual species that it is applied to. The RFLP technique is the most established molecular marker technique, however it is also time-consuming and requires relatively large amounts of DNA. The advantage of this technique is that there are well-established genetic maps in cotton based on RFLP markers, and so this simplifies the mapping of particular genes.

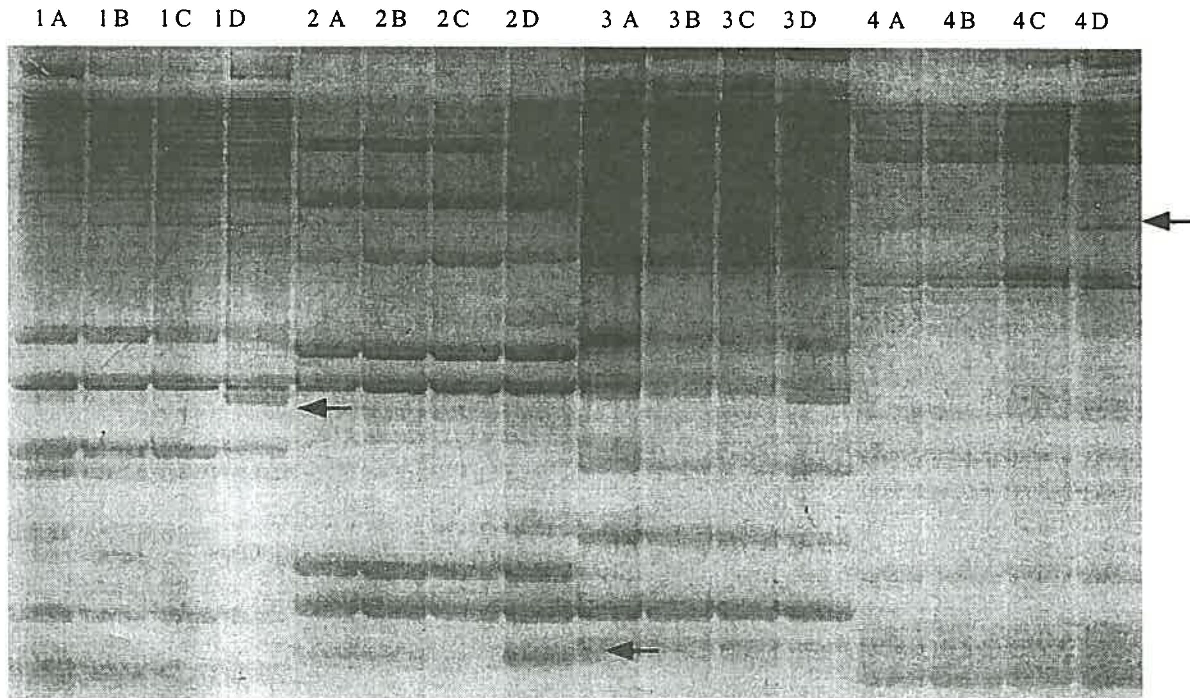


**Figure 1.** An example of an AFLP gel. A – resistant individuals, B – susceptible individuals, C – CS50 (resistant parent), D – Pima S-7 (susceptible parent). 1, 2, 3 and 4 are different PCR primer combinations used in the AFLP reaction. DNA fragments are separated according to size, and each band represents a discrete genomic region. Marker polymorphisms between CS50 and Pima S-7 are indicated.

## Molecular markers and *Xcm* resistance

This project utilised the AFLP and SSR techniques in conjunction with bulked segregant analysis to search for markers linked to the *Xcm* resistance locus in cotton. Bulked segregant analysis is a method for concentrating the search for markers to a particular region of the genome. Individuals of a segregating population are placed into bulks according to their phenotype. In this case, susceptible individuals would be placed into one bulk and resistant individuals into another. By comparing these two bulks, any marker differences between the two bulks should represent markers linked to the gene of interest. The degree of linkage will be determined by the size of the bulks, with larger bulks detecting only closely linked markers. In addition, previously mapped RFLP markers were used to establish the identity of the resistance locus present in Australian cotton cultivars.

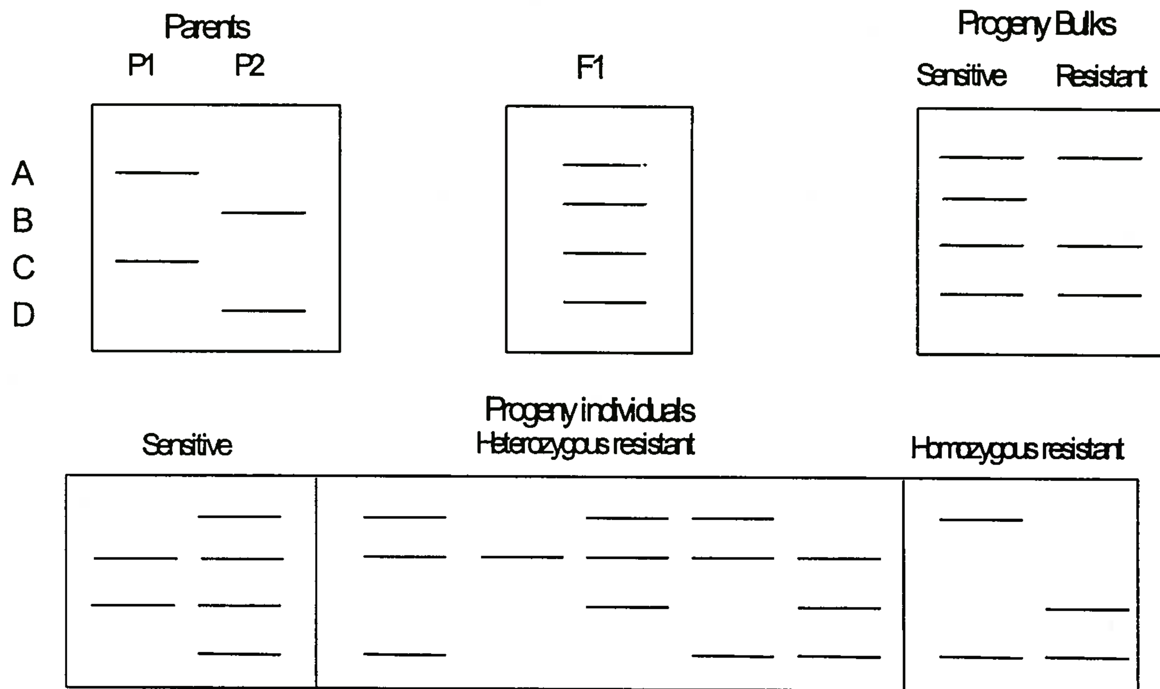


**Figure 2.** Schematic diagram of Bulked Segregant Analysis. By pooling sensitive and resistant individuals the search for linked markers can be concentrated to a particular region of the genome. If there are four markers (A, B, C, D) which show variation between two varieties (P1 and P2), then markers not linked to the resistance locus (A, C, D) will segregate independently in individual progeny and will appear in both resistant and sensitive bulks. If a marker is linked to the resistance locus, it will appear as a difference between resistant and sensitive bulks (B).

## The *Xcm* resistance locus

The resistance locus in Australian cotton cultivars segregates as a single locus, with resistance completely dominant, and the same locus is present in all the resistant Australian cultivars as it was derived from a single initial source. There are many different resistance genes, which give varying amounts of resistance, and different specific reactions to different races of *Xcm*. The most common *Xcm* race in Australia is Race 18, which is also the

most virulent of the more common races, and can overcome most of the known resistance genes, singly, and in combination. The reported source of the resistance genes from the USA is the $B_2B_3B_7$ resistance gene combination, which was introduced into the breeding program in the 1970's. However, the $B_2$ and $B_3$ genes have recently been mapped (Wright et al 1998) using RFLP probes, and while they map to the same chromosome, they are at opposite ends of the chromosome and should segregate relatively independently in a cross due to high levels of chromosome recombination between the two genes. In addition, the $B_2$ and $B_3$ genes reportedly do not confer to resistance to *Xcm* Race 18, either singly or in combination. It is therefore difficult to imagine how all three of these genes could have been maintained in the Australian cultivars or how they could segregate as a single locus. This suggests that the resistance present in our cultivars may have been derived from a different source than specified in their pedigrees. Our existing data, however, suggests that the resistance gene in the Australian varieties is the $B_{12}$ gene and not the $B_2B_3B_7$ gene complex. There is only one resistance gene that gives resistance to Race 18 by itself, and this is the $B_{12}$ gene. This gene was also mapped by Wright et al. The mapping population used to search for AFLP or SSR markers linked to the resistance gene was examined using the RFLP probes already mapped near to the $B_2$, $B_3$ and $B_{12}$ resistance genes respectively. The RFLP probe linked to the $B_{12}$ gene showed an extra band in the susceptible parent, and all 27 sensitive progeny examined so far showed the presence of this extra band. The resistant individuals in the mapping population were progeny tested to determine the zygosity of the resistance locus. In almost all cases, the presence of the extra band in resistant individuals was due to the fact that the individual was heterozygous for the resistance locus. From a total of 60 resistant plants tested, only 9 did not show segregation of the RFLP band according to the zygosity of the individual. When the probes linked to the $B_2$ and $B_3$ genes were used on the mapping population, they did not segregate according to resistance or susceptibility in the individual plants. Conversion of these RFLP probes to PCR-based markers will facilitate the processing of greater numbers of plants, and so improving the certainty of these conclusions. This will also allow for the search for alternative markers such as Single Nucleotide Polymorphisms (SNP's) in more closely related varieties (i.e. within *G. hirsutum*), where polymorphisms have previously been difficult to find.

## Mapping the *Xcm* locus

A total of 216 commercially available SSR markers were used to search for a marker linked to the *Xcm* resistance locus. While almost all were polymorphic between the two parental species, only one was polymorphic between the resistant and susceptible bulks. Further individuals were examined from the mapping population to determine the amount of linkage between the resistance locus and the SSR marker. From 31 susceptible plants examined, 8 showed recombination between the susceptible locus and the SSR marker, and from 21 homozygous resistant plants examined, 8 showed recombination between the resistance locus and the SSR marker. This equates to a recombination rate of approximately 30%, which indicates that there is a 30% chance of incorrectly scoring a plant as resistant or susceptible when using this SSR marker because it is not very close to the resistance gene. This commercially available set of SSR markers are too widely spaced to be very useful for fine-scale mapping, and to achieve modest resolution over 500 SSR markers scattered throughout the cotton genome would be required.

The AFLP technique has also been used to search for markers linked to the resistance locus. Approximately 3500 loci have been scanned. The rate of polymorphism between the two parental varieties is about 10%. To date no markers linked to the resistance locus have been found. This could be due to the fact that the AFLP technique is more sensitive than the SSR technique, and it can detect DNA molecules that form only a small proportion of the total population. This means that if in one individual from a bulk recombination has occurred between the gene of interest and a particular marker, then the bulks will be monomorphic at that locus and will not show a linked

marker. Consequently, a large number of loci must be scanned in order to find a tightly linked marker. Alternatively, the number of individuals in each bulk can be reduced. The fewer individuals in a bulk, the more likely a linked marker is to be found, but the degree of linkage will be reduced.

## Conclusion

In the search for markers linked to the *Xcm* resistance locus found in Australian cotton cultivars, one loosely linked SSR marker has been found to date. By itself, this marker is not particularly useful as it is so far from the resistance locus. However, if another marker can be found that is on the other side of the locus, then using these markers in conjunction may render the scoring of plants as resistant or susceptible by use of these markers more accurate. Alternatively, a more closely linked marker may be found using the AFLP technique. The use of previously mapped RFLP markers has facilitated the investigation of the identity of the resistance locus, which is the same in all resistant Australian varieties. The preliminary results of this investigation indicate that the resistance locus is different to that which is commonly believed to be present in Australian cultivars. Further investigation will elucidate these findings, and resolve the identity of the *Xcm* resistance locus found in Australian cotton varieties.

## Acknowledgments

This PhD project is funded by the Australian Cotton Cooperative Research Centre. We would like to thank Greg Constable and Peter Reid for providing the seeds for the varieties and crosses used in this project.

## Reference

Paterson, R.J., Thaxton, P.M., El-Zik, K.M., Paterson, A.H. (1998). D-subgenome bias of *Xcm* resistance genes in tetraploid *Gossypium* (cotton) suggests that polyploid formation has created novel avenues for evolution. Genetics 149 pp. 1987-1996.