

Genetic mapping in crop legumes (Objective 5, March 2007)

5A Provide novel germplasm for trait analysis (Fast Neutron mutant analysis and mapping)

5B Integrated Genetic Maps

Objective 5A will be achieved by

- a) selecting a set of 12 independent M₃ individuals from a novel legume genetic resource [fast neutron (FN) deletion pea population]
- b) performing cDNA-AFLP analysis of gene transcripts from total plants of these individuals and determining the number of genes deleted per FN line
- c) sequencing the cDNA-AFLP fragments corresponding to the missing transcripts and identifying the corresponding genes
- d) performing marker analysis of the deleted fragments by mapping identified genes in extant mapping RI populations

Objective 5B will involve

- e) identifying the most effective strategy for relating the genetic maps of pea, field bean and lupin for UK use
- f) establishing genetic mapping populations for bean and lupin within UK through coordination with the EU Grain Legumes project
- g) designing a comprehensive set of gene-based markers (at least 100) to enable integration of gene maps across crop species
- h) defining a set of molecular markers for priority traits that can be exploited by breeders

PROGRESS:

- a) selecting a set of 12 independent M₃ individuals from a novel legume genetic resource [fast neutron (FN) deletion pea population]

A set of FN mutant lines, including some with clear effects on plant architecture, have been selected for further analysis. In total, 16 M4 lines, including 8 FN lines that have no visible phenotype, were chosen, as listed below:

FN accession	Phenotype/Genotype
1216/9	Narrow leaf base (nlb)
2122/2	Stipules reduced (st)
1765/1	Apulvinii (apu)
2832/4	Apulvinii (apu)
1453/1	Arthritic (art1)
1579/1	Arthritic (art1)
1889/3	Creeping (creep)
2002/7	Creeping (creep)
2007/1	No mutant phenotype
2012/7	No mutant phenotype
2022/1	No mutant phenotype
2036/7	No mutant phenotype
2039/1	No mutant phenotype
2063/6	No mutant phenotype
2108/1	No mutant phenotype

In addition, two FN mutant lines affecting leaf form had been characterised independently by year 1 (*uni*, *cri*) and seven *tl* mutants were under investigation in year 1.

b) performing cDNA-AFLP analysis of gene transcripts from total plants of these individuals and determining the number of genes deleted per FN line and c) sequencing the cDNA-AFLP fragments corresponding to the missing transcripts and identifying the corresponding genes

In the year 1 report we described an AFLP approach to the analysis of FN mutant lines. We undertook this approach because we thought the transcript of interest was likely to be of low abundance, so in the present reporting period we compared this method with cDNA-AFLP. In the earlier approach we used bulked samples of DNA from independent FN alleles, so we do not know how many independent mutations there are per line, however we will assess this in the new set of selected lines.

The salient features of the cDNA-AFLP approach adopted are presented in Figure 5.1 below:

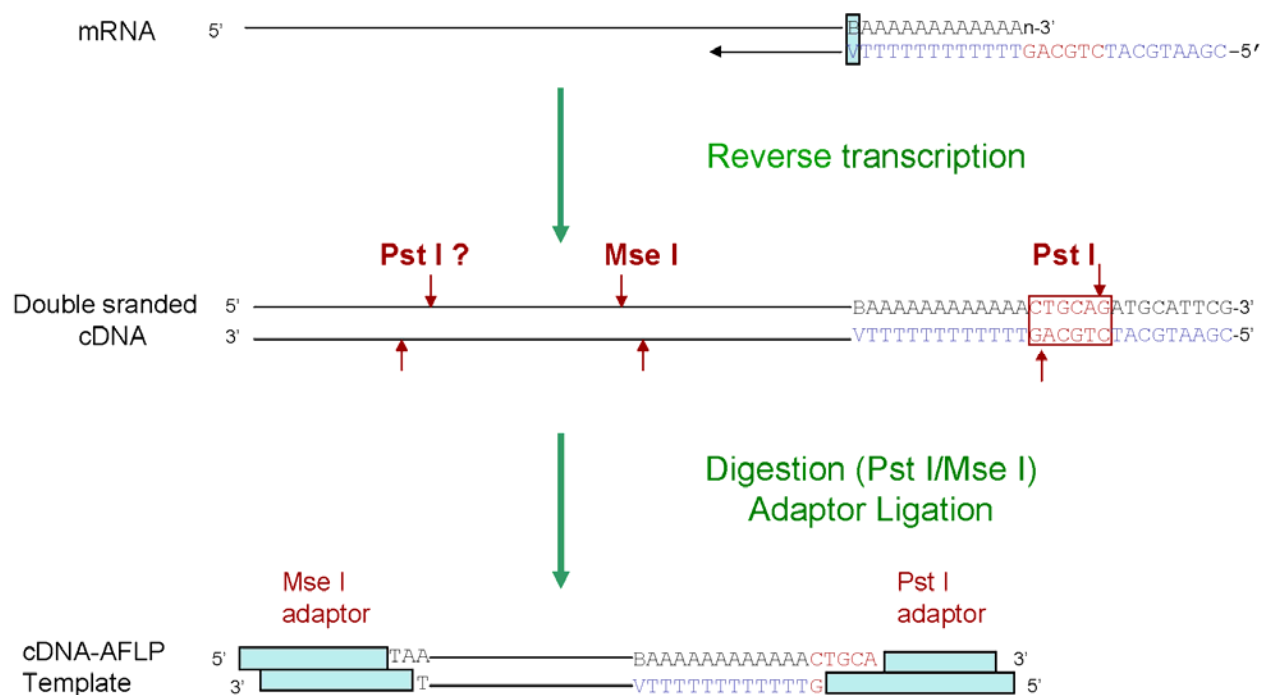


Figure 5.1: cDNA-AFLP strategy, where B is any base except A, V is any base except T

This procedure has several possible merits:

- 1) All amplicons correspond to genes so they can be positioned with reference to the *Medicago truncatula* reference sequence.
- 2) All cDNAs will carry a PstI site so all transcripts should be identifiable.

There are however some de-merits:

- 1) Some transcripts may be rare, and the method may detect only transcripts that are abundant or reasonably so.

- 2) The Pst primer is complex (Figure 5.2). It includes the A:T₁₂ tract and the selective bases are placed after this sequence:

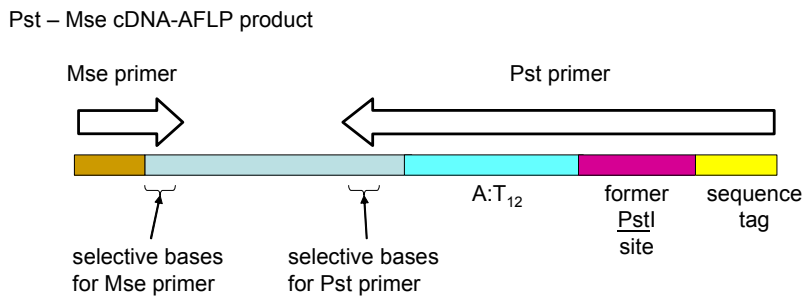


Figure 5.2

- 3) Because all transcripts are tagged with a 3' PstI site, the PstI site itself does not serve to reduce the complexity of the amplicons.

We therefore set out to test the seriousness of these problems, and decided to determine whether this method would detect the *crispa* deletion. We identified the MseI site nearest to the poly-A tail and used this to define the Mse selective bases. This involved determining the 3' end on the message, and there were several different poly-A sites that appeared to be used. One was selected for these experiments.

We needed to determine what degree of selection would be needed, and so tested a range of amplification strategies.

We reasoned that all transcripts would contain a PstI site, by design, and that all would also carry an MseI site within amplification range of the 3' terminus of the transcript³. If not, we would miss the transcript and MseI sites are very common especially in AT rich sequences such as 3'UTRs. If we further estimate there are about 30,000 genes in pea (by analogy to the estimate from *M. truncatula*), then we want to see about 30 amplicons per selective set. We need $\log_4(1,000)$ selective bases, suggesting that about 5 selective bases are needed.

Figure 5.3 below, based on tests using the *crispa* mutation, shows the result of amplification with several different degrees of selection.

³ Note: An MseI site (TTAA) could precisely overlap the beginning of the poly A tail and all such transcripts would be missed by this procedure.

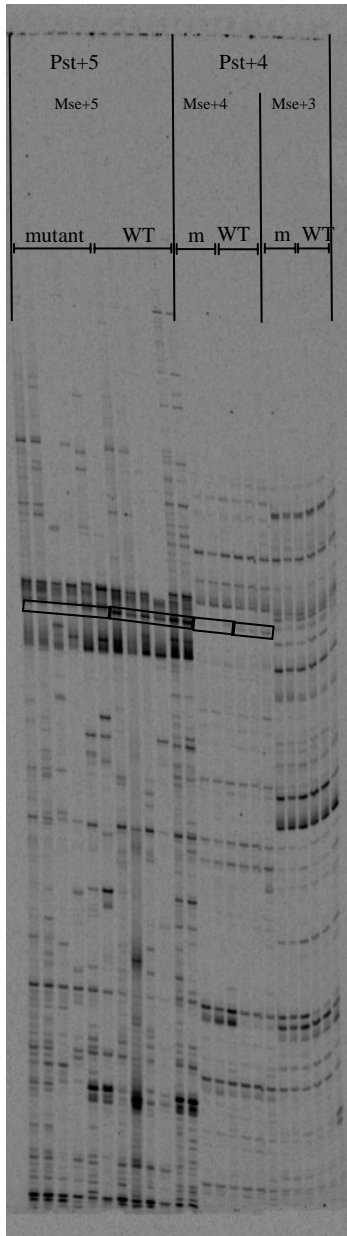


Figure 5.3

This suggested an amplification protocol with 5 selective bases on each of the Mse and Pst primers, requiring a four-step amplification procedure: Mse+a, Pst+b followed by Mse+X, Pst+Y

Our general conclusion from these experiments is that the cDNA-AFLP approach had the advantage that it may give detectable results where deletions are small and the resulting defective message has a rapid turnover. However the method is complex and the resulting banding patterns have a wider range of band intensities compared to AFLP with a DNA template, and together these make band calling unreliable.

It appears from our analysis of ten FN deletions that all have deleted segments of DNA comparable in size to a gene. This mean that the chance of a deletion including a Pst site is reasonably good, so the anticipated difficulties with the direct AFLP approach are less of a problem than expected. On balance the direct AFLP approach is the better of the two methods in this system and will be used in the future. The AFLP methodology has been used successfully for the identification of the *tI* gene (see last year's report).

d) performing marker analysis of the deleted fragments by mapping identified genes in extant mapping RI populations

All of the deleted sequences isolated so far correspond to known, mapped genes.

Objective 5B will involve

e) identifying the most effective strategy for relating the genetic maps of pea, field bean and lupin for UK use

This has been done through the management group and stakeholder consultation.

The demand for lupin maps is low, and there are publicly available lupin maps, and the relationship between these and the *M. truncatula* genome sequence is being determined in the EU FP6 Integrated Project 'Grain Legumes' and also by labs in Australia that collaborate with the corresponding EU partner lab. As there are no specific UK lupin mapping populations (or interest in generating these) we are of the view that lupin genetic needs are best served by the international community and that generating these populations and maps is not an effective use of PCGIN resources.

For faba bean, genetic maps exist and are being developed in the EU (Cordoba) and Australia (Perth), and mapping populations are shared by these labs (through the EU FP6 project). Thus there is some activity in faba bean mapping that could be drawn upon. However, these maps are not between crosses of UK cultivars and do not fulfil the needs of both the winter and spring types. Therefore we have initiated a collection of faba bean lines with a view to extracting suitable inbred lines that can be used for the generation of mapping populations. This calls on PCGIN to establish such a UK collection, undertake some genotyping to assess the pattern and extent of genetic diversity within these lines and to extract inbreds prior to generating segregating populations. The NIAB lab has taken the lead on this initiative with the generation of a 'white paper' on faba bean genetic objectives for the UK to be published on the PCGIN web site. JIC and NIAB with several collaborators have also submitted a grant proposal to the BBSRC-Dfid initiative that, if funded, would go some way to the establishment of a faba bean genetics programme in the UK. Following consultation and discussion, lines of *Vicia faba*, together with *Vicia narbonensis*, have been grown in year 2 under greenhouse conditions, and seeds distributed for phenotyping at three sites in year 3 (See Objective 2).

For pea, multiple genetic maps are available, and these are now well connected to one another through common markers and by being anchored to the *M. truncatula* genome sequence. The availability of maps *per se* is not problematic, but some investment in the interrelation of these maps with informatic tools is needed. The development of the PCGIN mapping populations discussed in Objective 4 will provide the mechanism for relating the determinants of agronomic traits of interest in the UK to these maps.

f) establishing genetic mapping populations for bean and lupin within UK through coordination with the EU Grain Legumes project

See (e) above.

g) designing a comprehensive set of gene-based markers (at least 100) to enable integration of gene maps across crop species

Approximately 1000 genes have been used for the generation of intron-directed PCR markers in the lab of Dr György Kiss at Gödöllő in Hungary, within the framework of the EU FP6 project. (This information is publically available at: <http://bioweb.abc.hu/mt/pisprim>). In collaboration with JIC, ~300 of these have been mapped in a JIC mapping population. JIC is currently involved in further genetic mapping with these markers with the aim of refining and selecting appropriate sets of markers for genetic mapping.

Integration of genetic maps of pea, whether with gene specific markers or not, will also

serve to anchor maps of new populations to the *M. truncatula* gene space sequence. The provision of simple, cheap marker systems for anchoring pea maps is thus an additional objective.

h) defining a set of molecular markers for priority traits that can be exploited by breeders

This activity is pending the identification of the relevant traits and segregating populations. Discussions have been held with the industry in relation to the transfer of markers linked to agronomic traits and their ease of use and reliability (Nickerson-Advanta and Unilever).