General enquiries on this form should be made to:
Defra, Science Directorate, Management Support and Finance Team,
Telephone No. 020 7238 1612
E-mail: research.competitions@defra.gsi.gov.uk

SID 5 Research Project Final Report

- **Note**
  In line with the Freedom of Information Act 2000, Defra aims to place the results of its completed research projects in the public domain wherever possible. The SID 5 (Research Project Final Report) is designed to capture the information on the results and outputs of Defra-funded research in a format that is easily publishable through the Defra website. A SID 5 must be completed for all projects.
- This form is in Word format and the boxes may be expanded or reduced, as appropriate.
- **ACCESS TO INFORMATION**
  The information collected on this form will be stored electronically and may be sent to any part of Defra, or to individual researchers or organisations outside Defra for the purposes of reviewing the project. Defra may also disclose the information to any outside organisation acting as an agent authorised by Defra to process final research reports on its behalf. Defra intends to publish this form on its website, unless there are strong reasons not to, which fully comply with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.
  Defra may be required to release information, including personal data and commercial information, on request under the Environmental Information Regulations or the Freedom of Information Act 2000. However, Defra will not permit any unwarranted breach of confidentiality or act in contravention of its obligations under the Data Protection Act 1998. Defra or its appointed agents may use the name, address or other details on your form to contact you in connection with occasional customer research aimed at improving the processes through which Defra works with its contractors.

Project identification

1. Defra Project code: AR0711

2. Project title:
   Pulse Crop Genetic Improvement Network

3. Contractor organisation(s):
   John Innes Centre
   Norwich Research Park
   Colney
   Norwich NR4 7UH

   NIAB
   Huntingdon Road
   Cambridge CB3 0LE

   PGRO Research Station
   Great North Road
   Thornhaugh
   Peterborough PE8 6HJ

4. Total Defra project costs (agreed fixed price): £ £1,213,153

5. Project: start date ..... 01 February 2005

   end date ..... 01 May 2009

6. It is Defra's intention to publish this form.
Please confirm your agreement to do so.....**YES** ☒ **NO** ☐

(a) When preparing SID 5s contractors should bear in mind that Defra intends that they be made public. They should be written in a clear and concise manner and represent a full account of the research project which someone not closely associated with the project can follow.
Defra recognises that in a small minority of cases there may be information, such as intellectual property or commercially confidential data, used in or generated by the research project, which should not be disclosed. In these cases, such information should be detailed in a separate annex (not to be published) so that the SID 5 can be placed in the public domain. Where it is impossible to complete the Final Report without including references to any sensitive or confidential data, the information should be included and section (b) completed. NB: only in exceptional circumstances will Defra expect contractors to give a "No" answer.
In all cases, reasons for withholding information must be fully in line with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.

(b) If you have answered NO, please explain why the Final report should not be released into public domain

Executive Summary

7. The executive summary must not exceed 2 sides in total of A4 and should be understandable to the intelligent non-scientist. It should cover the main objectives, methods and findings of the research, together with any other significant events and options for new work.

The broad objective of this network was to align basic research in legume genetics with the needs of growers and users. The overall strategy was to undertake genetical research that could impact on breeding.

The project has focussed on two species, pea (*Pisum sativum*) and faba bean (*Vicia faba*). These represent several crop types ranging from the production of dry seed for animal feed to fresh vegetables for human consumption. The two species are closely related, but have different breeding systems, pea being inbred and faba bean open pollinated. Thus the genetics of the two is very different. The pea genome is large, about the same size as the barley genome, while faba bean, although a diploid, is about the same size as the wheat genome. For these reasons our main focus has been on pea as the more tractable species.

The network aimed to integrate activities in diverse locations and this was achieved through a series of open meetings that ranged from general dissemination activities to those focussed on specific problems, notably the emergence of bruchid pests as a major constraint on the value of UK faba bean production. We have developed a web-site that allows us to easily communicate general and specific project information, and run a database of stakeholders and interested participants.

We recognised that much work in pea has been done on the one hand studying the genetics of exotic lines and on the other the performance of cultivars. We sought to redress this in two ways. The first was to include genotyped exotic lines in field experiments at several locations, and the other was to generate genetic information on the relationship between cultivars for the generation of populations for the study of trait genetics.

These two approaches have had interesting outcomes. We have identified a source of resistance to downy mildew in pea and have identified a possible source of enhanced standing ability among the exotic lines. These two traits will be studied genetically with a view to providing genotypes that can be used in the initial stages of cultivar breeding. In consultation with stakeholders, we have also trialled a range of bean genotypes from Australia, alongside UK varieties, and assessed their performance. Among the Australian lines, there were lines that ranked well and these may be exploited for breeding for a drier UK climate.

In a complementary approach we have identified suitable parents for the generation of segregating populations in pea that can be used for the genetic dissection of traits that are the targets of breeders. The

generation of three inter-cultivar recombinant inbred populations is a significant achievement that provides valuable genetic materials for the dissection of traits under field conditions.

For basic genetics we wanted to find novel sources of variation and to characterise the genetic basis of defined phenotypic characteristics. These objectives led us to the creation and analysis of systematic mutant pea populations, enabling the isolation of genes known only by their phenotypic effect together with the identification of genetic variants in selected target traits.

All of these activities have progressed along with the refinement of genetic maps and the generation of molecular markers. We have been able to align the pea genetic map with the sequence of the genome of the model legume *Medicago truncatula* and have generated gene specific markers for faba bean, showing that cultivars are internally heterogeneous.

For seed quality, we have isolated novel mutants for antinutritional proteins that impact on animal feed quality. Two classes of proteins were targeted, one mutant being a naturally occurring mutation in exotic germplasm, whereas mutants for the second were isolated from a mutagenised population of seeds that provides a resource for novel mutations.

For food quality, stakeholder interest defined two areas in which scientific research was required on pea seeds. The first of these was seed bleaching, a phenomenon that can reduce the value of the crop significantly and, in extreme situations, result in the value of the crop being equivalent to that of animal feed.

We have identified, among exotic germplasm, lines with extremely stable green colour, one of which has been used to establish a population of recombinant inbred lines, using a commercial bleach-prone line as the second parent. These lines have been used to make a genetic map that will inform future studies of quantitative variation in seed bleaching under field conditions.

Besides this 'open' approach to gene identification, we have studied variation in two candidate genes, either or both of which may influence the pathway of colour degradation. Genetic markers have been devised for these genes, which can be used by the industry in breeding programmes. Two variants have been identified for one of these genes, resulting in both cases in a reduction of an enzyme activity potentially linked to colour loss. These variant genes are being introduced into commercial lines. A first trial of the lines generated for one of these genes (backcrossed once to commercial parent, following initial cross) showed that there was a greater retention of colour by the lines carrying the variant gene.

The second area defined by stakeholders involved the definition of seed quality for food use. In joint experiments with CSL (FERA), York, we have explored the potential for defining metabolites or small compounds that may be linked to taste, flavour and health. Preliminary experiments have shown that this approach can be linked to the genetics of seed quality, and this has been pursued through the development of a LINK project proposal, involving four research partners together with seven industrial partners, providing nine sources of in kind contribution to the project.

Project Report to Defra

8. As a guide this report should be no longer than 20 sides of A4. This report is to provide Defra with details of the outputs of the research project for internal purposes; to meet the terms of the contract; and to allow Defra to publish details of the outputs to meet Environmental Information Regulation or Freedom of Information obligations. This short report to Defra does not preclude contractors from also seeking to publish a full, formal scientific report/paper in an appropriate scientific or other journal/publication. Indeed, Defra actively encourages such publications as part of the contract terms. The report to Defra should include:
   - the scientific objectives as set out in the contract;
   - the extent to which the objectives set out in the contract have been met;
   - details of methods used and the results obtained, including statistical analysis (if appropriate);
   - a discussion of the results and their reliability;
   - the main implications of the findings;
   - possible future work; and
   - any action resulting from the research (e.g. IP, Knowledge Transfer).

The Pulse Crop Genetic Improvement Network (PCGIN) 2005 -2009

The activities planned for the PCGIN 2005 – 2009 were organised into six primary objectives. This report documents the progress achieved within these objectives as separate entities, although in practice there is much overlap and interaction between their component parts.

Summary of objectives and sub-objectives:

Objective 1: Communication and Delivery

To establish and promote effective communication between the major players responsible for the genetic improvement of pulse crops. This first objective is not a scientific activity *per se* but it is essential to the relevance, application and delivery of the scientific activities of the PCGIN. Its activities are to

a) establish and promote good communication between UK breeders in the pulse crops and the research base, both basic and applied, and ensure all have the opportunity to provide feedback throughout the project
b) establish a close liaison with SASA to ensure coordination and integration of our activities
c) establish a database of ongoing research in the area of pulse crops especially in the UK, but also at an international level
d) establish a web site, updated every 4 months and linked to the other Defra crop improvement networks
e) enable interaction days for management and stakeholders groups, providing opportunities to view and discuss genetic material
f) develop associated projects linking participants

Objective 2: Phenotyping

Phenotypic characterisation of novel legume germplasm

This will involve

a) evaluating priority traits for different species by members of the PCGIN, taking account of the assessment of priorities already indicated by the breeders' survey
b) making informed choices on the species, germplasm, traits and specific biotic stresses to be studied, to include the use of exotic germplasm
c) developing a series of protocols for assessment of the traits required, with rapid and reproducible scoring techniques and reliable methodology for G x E interactions established at three sites (NIAB, JIC, PGRO)
d) correlating genetic marker data, where available, with phenotypic characters

Objective 3: Performance

Extant varietal performance data will be associated with genotype data

This will be based on

a) selection of maximally informative databases for UK pulse crop characters, site characteristics, history and yearly records
b) selecting a set of 50 cultivars that are differential for the traits identified in the assessment of breeders' priorities (yield, standing ability, disease resistance), based on maximally informative databases
c) genotyping selected cultivars with genetic markers (200 each)
d) selecting three cultivars that are maximally informative on the basis of phenotypic data and marker analysis for the generation of segregating populations
e) establishing recombinant inbred lines (RILs) to F_6 from crosses between the chosen lines (150 lines per population).

Objective 4: Reverse genetics

TILLING *for genes that regulate the development of the aerial part of the legume plant*

This will involve

a) access to ca. 5000 pea cDNA sequences (ESTs).
b) annotating sequences and identifying targets for TILLING, based on database mining and identity of orthologous genes involved in plant shoot architecture
c) designing primers based on candidate genes, derived from a) and b) and from databases developed within the EU project (Grain Legumes)
d) TILLING for mutants in these genes, using the platform developed within the EU project (Grain Legumes) and assessing phenotypes of mutants.

Objective 5: Genetic mapping in crop legumes

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 M_3 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

Objective 6: Genetics of Seed Quality Traits

This will involve
a) expansion of Objective 1 to include industrial end users within the range of interested parties
b) wide consultation with a wide range of industries to define priority seed quality traits
c) development of tools and definition of protocols for analysis of these traits
d) establishment and analysis of recombinant inbred populations that segregate for key traits in relation to seed and protein quality, as defined by these end-users
e) *de novo* satellite projects with industry to exploit this variation

Report of achievements within objectives and sub-objectives:

Objective 1: Communication and Delivery

To establish and promote effective communication between the major players responsible for the genetic improvement of pulse crops.

We have established a network for effective communication among breeders, end users and the scientific research base. This is evidenced by the web site (www.pcgin.org) which has provided a vehicle for interested parties to record their specific interests in pulse crop research. There are currently 75 registered members on the PCGIN web site.

The web site (www.pcgin.org) provides links to the other Defra-funded crop improvement networks: Wheat Genetic Improvement Network (WGIN: www.wgin.org.uk); Oilseed Rape Genetic Improvement Network (OREGIN: www.oregin.info) and Biomass for Energy Genetic Improvement Network (BEGIN: www.biomass4energy.org). It is updated regularly, being used primarily for the advertisement of relevant meetings and events, but also for dissemination of information relating to genetic maps and associated sequence data. Registered members can access certain preliminary documents within the password-protected areas, allowing for input and feedback based on data prior to publication in journals. The web site provides a database for email dissemination of information and alerts to relevant meetings, newspaper articles.

The web site (www.pcgin.org) has 5 main sections:
1. Project Information
2. View Project Data
3. Login as PCGIN USER
4. Register as user / member
5. General Interest Articles

The fist of these allows navigation through progress within the project objectives and provides access to the various data files relating primarily to plot trial and phenotyping work. Relevant excerpts from the annual reports are also available here. The second of the five sections gives access to genetic information that has been generated at national and international level, and includes a link to the Legume Information System (LIS).

In every year covered by this report, formal management meetings were held regularly to discuss organisational aspects of the project. Stakeholder meetings have been held annually, either presenting the project in a general way or with a specific focus (for example, a focussed meeting on seed quality in October 2007). The agendas and minutes of all meetings, once approved, are posted on the web site. More informal meetings to view and discuss the plots at the three locations (PGRO, NIAB and JIC), and to view plant material in greenhouses, also took place annually. This included participation by stakeholders, for example, Sharps/Nickerson-Advanta and Unilever/Birds Eye. Specific visits by industry have been hosted throughout by the three scientific organisations; for example,

representatives form Heinz visited JIC on two occasions (December 2007 and January 2008), which resulted in participation by this company in the stakeholder meeting in December 2008.

PCGIN has developed and maintained close contact with SASA, with the latter attending several of the project meetings, and all the stakeholder meetings. In particular, SASA was involved in the choosing of materials for genotyping and obtaining the relevant phenotypic data pertaining to these lines. Additionally, PCGIN has assisted SASA with the identification of seed proteins that are being used as diagnostic markers for profiling genotpyes (Tom Christie, SASA, personal communication).

The PCGIN has been promoted through poster presentations, hand-outs and displays at various annual local and national events (for example, PGRO members' day, Cereals exhibition), as well as at various international conferences (for example, the recent IVth International Conference on Legume Genetics and Genomics, Puerto Vallata, Mexico, December 2008). Seed collections representing diversity and variation in genes under investigation were used, in conjunction with a PCGIN postcard and plant displays, as demonstration material at various meetings and conferences and to promote discussion at open events with diverse audiences, for example the annual PGRO members' day. PCGIN research contributed to invited lectures and talks, presented to several national and international conferences and also to industry. These have been documented in the annual reports.

The PCGIN was linked formally to EU-GLIP (www.eugrainlegumes.org) through joining the Technology Transfer Platform (GL-TTP). PCGIN research contributed to both of the GL-TTP workshops that have been held so far: 'Targeting Science to Real Needs', Paris in April 2007 and 'Integrating legume science and crop breeding' in Novi Sad, Serbia in November 2008. The latter GL-TTP workshop took a strategic decision concerning the realignment of the platform and the European Grain Legumes Association (AEP) with the aim of merging these two organisations. PCGIN will remain affiliated to these organisations, in whatever merged format they take.

Associated projects have been established that link participants. One of these was a PCGIN satellite project, part funded by the EU (GLIP), that was conducted by CSL, York in collaboration with JIC and exploited metabolite profiling to define climatic changes in plant composition. A CASE studentship was set up between JIC and Unilever and was supervised by both organisations to build a comprehensive genetic linkage map for a wide cross in pea (David Holden, PhD 2009). When the bruchid beetle (*Bruchus rufimanus)* was acknowledged by the industry as a growing problem for *Vicia faba* (faba and field beans) in the UK, a student project was set up (M.Sc. in Plant Genetics and Crop Improvement, JIC-UEA, Norwich) to explore natural variation in defensin genes in pea, to include *P. fulvum* lines found to have resistance to *Bruchus pisorum.*

An employee of Limagrain UK is currently registered as a M.Sc. student (joint JIC-UEA course in Plant Genetics and Crop Improvement), and this studentship includes a research project linked to PCGIN and industrial objectives. A number of LINK projects that are connected directly or indirectly with PCGIN have been established: the Bruchid and GreenPig projects both involve NIAB and PGRO. Based on PCGIN Objective 6, the PCGIN satellite project, and extensive industrial discussion, a LINK pre-proposal was developed. This was accepted and developed to a full proposal 'Understanding Quality Determinants in Pea Seeds to improve market opportunities that promote sustainable agriculture (QDiPS: drivers of a sustainable agriculture)' that was submitted in March 2009 and presented to the programme panel committee at their meeting (31/3/09). This project has been recommended for funding.

Objective 2: Phenotyping

Phenotypic characterisation of novel legume germplasm

A questionnaire among UK breeders of the research requirements and common limitations to the major legume crops in the UK was undertaken in advance of the project formulation, the results of which were used to inform and shape the work plan that was subsequently undertaken. The major issues identified priorities for all the legume crops surveyed (combined and vining peas, field beans and lupins) and included factors controlling yield, yield stability and disease and pest resistance. These were broken down for the individual crops and among the highest priority areas were factors affecting plant architecture in peas (both combined and vining) which equated to standing ability, and drought tolerance/ sensitivity for faba beans.

Genetic resources available within breeding programmes represent a limited subset of the broader variation present within each given species. While many breeders recognise the value of 'exotic' germplasm as a source of novel traits and allelic variation, many lack the capacity to evaluate wider germplasm that could be of use in their programmes. Prospecting exotic germplasm to identify germplasm with potentially useful traits that offer long-term benefits, while also extending the gene pool and resources used in breeding programmes, was widely endorsed as an objective for the programme. We therefore undertook to step back from current breeding material and evaluate a range of exotic cultivated germplasm under microplot conditions across the sites of the main project partners (PGRO, NIAB and JIC) on the basis of agreed protocols. These trials were conducted over three years of the project to enable elements of stability of observed responses to be assessed. The three sites are of contrasting soil type, PGRO being a silty loam, NIAB a loam and JIC a free draining sandy loam and all are

suitable for peas and beans. These sites thus offered a good test of stability of performance while also enabling members of the management team, technical support workers, breeders and other stakeholders the opportunity to observe and discuss a wider range of variation within these species than was normally encountered. The three growing seasons of these trials were dramatically different. Year 1 saw the weather change in early June as plots were coming into flower, to become very hot and dry. These conditions persisted to give one of the shorted flowering to maturity seasons in many years. Year 2 was overall a very cool wet summer which brought about particular disease and harvesting problems. Year three was the most 'average' season of the three being of moderate temperatures and regular rains which saw longer vegetative growth.

Pea was the main focus of this objective and the maximum capacity of microplots that could be managed by the three sites was set at 60. Three replicate plots of any one line at each site dictated that a maximum of 20 lines would be grown each year at each of the three sites. For ease of management, all trials were conducted on spring types. The results from the first year of multi-site microplot trials were analysed and a number of the less promising lines dropped in order to allow other material to be included. A total of 9 pea lines from the original 20 were taken forward after the first year together with three new accessions. 8 bean varieties were included for years 2 and 3. Details of the lines used and the results are presented in **Annex 2.1**; the general conclusions are summarised below:

A: Peas

Seed multiplication and selection of germplasm

Before the official start of the project, it was evident that the availability of exotic germplasm would be dependent on seed availability. Multiplication of 47 diverse germplasm lines (*Pisum sativum*) was carried out on a single site as single unreplicated microplots (Table S2.1a.1 in Annex 2.1). The selection of 27 germplasm accessions for this exercise was based on a phylogenetic analysis of marker data (SSAPs) from an extended set of cultivated *P. sativum* accessions (Vershinin et al, 2003). Five further lines (JI 15, JI 399, JI 281, JI 813, JI 1194) were included on the basis of their being parents of public mapping populations developed at the JIC, the data and material of some of which have been utilised within this project (*Objective 5B: Integrated genetic maps*). One further line (JI 2822, parent line for the JIC Fast Neutron mutagenesis programme) was selected on the basis of its additional uses within in the project (*Objective 5A: Fast neutron mutant analysis and mapping*). Together with the 18 germplasm accessions, two current commercial cultivars Bilbo and Cooper were included in year 1 of multi-site trials (2006) as standards against which the exotic material could be compared.

Following the multi-site microplot trials in year 1, the number of pea accessions taken forward into years 2 and 3 was reduced to 9 in order to allow further germplasm to be included. The three new accessions of pea that were introduced were three recombinant inbred lines (JI 2947, JI 2959 and JI 3142) that were developed as part of one JIC mapping population (JI 15 x JI 1194), which had previously been scored as potentially of improved standing ability.

Lodging and Crop Attitude

The introduction of the *afila* gene (*af*) in the mid 1970's, which results in additional tendrils replacing leaflets, was responsible for a significant improvement in the standing ability of the crop (Snoad 1974, Hedley and Ambrose 1981). The character was quickly adopted by pea breeders around the world and the majority of new cultivars released today incorporate the afila character, albeit in semi-leafless form (*afila* but with normal leafy stipules). Despite these improvements, standing ability of the pea crop is still a problem and is associated with difficulties in harvesting crops and, over years, stability was identified as a high priority trait by breeders where additional input and resources were required. The growing of microplots of exotic germplasm (all normal leafed types) allowed assessment for lodging at maturity across different sites and years. Two current commercial varieties were included as controls (Bilbo and Cooper). Cooper, was maintained in the trials over all three years. Despite the highly contrasting growing seasons of the three years, the first point to note is the striking difference between sites with respect to lodging, although close examination of the data found underlying trends (Annex 2.1).

Of all the exotic lines scored with some replicates not lodging, JI 2201 was scored most consistently as not lodged at JIC over the three years (Fig. 2.1). This line was picked out in the first year as having particularly wiry stems of moderate thickness (4mm). The same line was only recorded as not lodged in one replicate in one year at NIAB. Although it did not prevent lodging, the same wiry stem feature was noted in the material in all three sites. Clearly this observation warrants further study firstly to establish the attributes of the stem and to try to follow the character through segregation while also trying to introduce it into more commercial backgrounds. Furthermore, developing near-isogenic lines for this trait in combination with *afila* would allow better evaluation of its potential to further improve standing ability in pea.

The clear ranking of sites with respect to the lodging scores is an interesting finding. The basis of the differences between sites is not known and could be due to a combination of soil type and interaction with climate. Whatever the reason, the ranking of sites is one that could be used in the future when testing other material for standing ability.

Figure 2.1: Plots of microplots and mid-podding stage in 2006 of a. JI 2201 still standing well and b. JI 1194 severely lodged.

Conclusion of pea multi-site microplot trials

- Exotic germplasm selected on the basis of genotype data across the spectrum of diversity of *Pisum sativum* showed a range of highly contrasting plant architecture and canopy structures.
- The semi-leafless plant model proved the more resistant to lodging across sites and years than conventional leafed model
- The three sites ranked in terms of the number of lines resisting lodging were in the order: JIC NIAB and PGRO. These sites therefore offer a graded series that could be used further to quantify lodging resistance.
- An indication of improved standing ability was identified in JI 2201.
- The best performers among the exotic lines for seed weight across the trials were JI 1194 and JI 813.
- Higher biomass of JI 2201 under cool wet conditions suggested this form might offer potential as a fodder crop in certain regions such as the north and western regions.

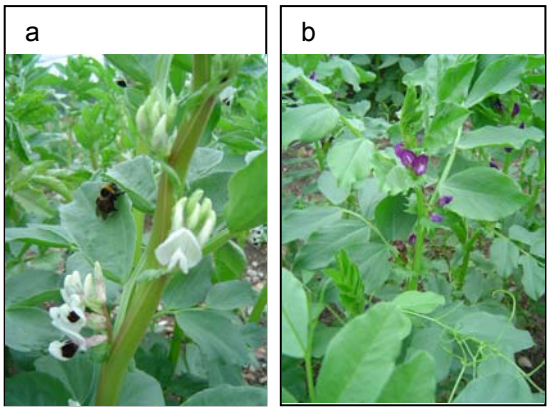
B: Field Beans

Phenotypic work on field beans was initially restricted to observations made on germplasm accessions grown out or regenerated from the *Vicia* collection maintained at JIC. In year 1, 60 accessions were grown as unreplicated microplots using the experimental plots at the JIC. Project partners and stakeholders were able to view and comment on this material which was all of European origin. In addition, regeneration of accessions under glass where individual plants are bagged to eliminate the risk of outcrossing also provided the opportunity to record phenotypic data on accessions.

Soil moisture has long been known to be one of the major limiting factors to yield of the faba bean crop (Lawes et al. 1983) and was identified as being of high priority in the assessments made by UK breeders as part of the development of this programme. Sensitivity or tolerance to drought is a complex series of traits and, during discussions in both management and stakeholder meetings, it was evident that aspect of components of yield and how they contributed to overall yield were poorly understood. It was agreed in year 1 to source novel germplasm from outside Europe that might exhibit adaptive variation that could be evaluated within the project. Such material, by its very nature, would be of different habit and would thus contrast to UK adapted material. Novel field bean germplasm was grown as part of the multi-site microplots trials in years 2 and 3, replacing of some of the less promising pea lines following trials in year 1. This material was then monitored for agreed traits and data collected from harvested individual plants from the inner part and the outside edge of every plot.

Germplasm was sourced from Australia where the field bean crop is well established and there is a similar interest in quality export markets to those targeted by the UK. The recommended variety literature for this material includes assessments of varietal performance by rainfall zone (Egan and Crouch 2009). A total of 5 varieties on the current recommended list in 2006 from Southern Australia were sourced. Alongside these, Fuego, the top selling UK spring field bean and Maris Bead, originally released in 1964 but of current interest due to evidence of possible disease resistance, were included. The latter is a small seeded tic variety that is grown and has a highly contrasting plant habit compared with Fuego which is a larger pale seeded variety, developed for the export market. The eighth *Vicia* line grown out in the trials was a narbon bean (*Vicia narbonensis*) cultivar 'Tanami' and represents a different taxa to that of field bean (Vicia faba). One of a number of *Vicia* taxa that are thought to be possible progenitors of faba bean, this one has never been established successfully. *V. narbonensis* and *V. faba* have never been successfully crossed to yield viable embryos. The line was included, firstly because it offered such a contrast to the faba types with conspicuous terminal tendrils and racemes of 1-2 flowers (Fig. 2.2b) and, secondly, Tanami takes its name from the Tanami desert and is a reference to the development of this species as a crop in low rainfall areas (Siddique et al 1996). The full list of Vicia germplasm sourced for the multi-site microplot trials is presented in Annex 2.1.

Figure 2.2: Photographs showing the racemes of a. Maris Bead (*V. faba*) and b. Tanami (*V. narbonensis).* In addition note the thicker stem of Maris Bead and the presence of conspicuous tendrils in Tanami.



The floral racemes of faba beans have anywhere between 2 and 9 flowers at a particular node. While this represents a very high yield potential, there is frequently significant loss of flowers at increasing nodes and under stressed conditions. Flower number and auto fertility was examined and results are summarised in Annex 2.1.

Faba beans have stronger stem bases than peas and are generally much more resistant to lodging. However, the crops are often left standing while other crops are harvested, so standing ability at harvest is an important character and forms part of the evaluation of new varieties as part of the national list process. Lodging scores varied across sites and years and are reported in Annex 2.1.

Conclusions on bean multi-site microplot trials

- A range of contrasting forms of faba was sourced that was novel to the UK and therefore of value for evaluation.
- Very high seasonal variation demonstrated the sensitivity of the crop to restricted rainfall in the establishment phase of the crop.
- Seed weight per plant data showed a strong Genotype X Environment interaction when regressed against environmental means. This was more pronounced in the faba lines than in the narbon line which had a much lower slope.
- Of the faba lines, UK material performed better than the Australian lines. Fuego, the most recent of these, outperformed all other lines overall for seed yield and in the ratio of seed to pod.
- Of the Australian lines, Icarus recorded higher seed yields from the inner plants in two trials, ranking 2nd overall. Cairo ranked 1st overall in one such trial, and 3rd in three trials. Both Icarus and Fiesta performed well for outer plant seed yield (1st or 2nd overall) in two trials.
- The narbon cultivar Tanami was a significant contrast to all the faba material and was of general interest in discussions. The 5th lowest intercept of seed weight regressed against the environmental mean for inner plants supports the claims for its cultivation in poorer environments and possibly in areas of low rainfall.

Objective 2: Disease testing

The aim of this part of the project was to add disease resistance data to the other characters being evaluated in microplot field work with exotic *Pisum* lines. The relative tolerance of pea lines to a range of diseases is an important consideration for farmers in relation to the decision of whether or how often to include a pea crop in a rotation. Accordingly we undertook both glasshouse and field experiments to assess the extent of resistance or tolerance of pea varieties to downy mildew (*Peronospora viciae* f.sp. *pisi*), powdery mildew (*Erysiphe pisi*), *Fusarium solani* f sp *pisi, Phoma medicaginis* var. *pinodella* and *Aphanomyces euteiches.* In addition to these studies an associated CASE PhD studentship between Unilever and JIC (Holden 2009) undertook a study of the genetics of resistance to downy mildew in some work undertaken at NIAB (the thesis is available on request of from the JIC library). The results of these investigations at NIAB and PGRO are available in Annexes 2.2 and 2.3, respectively.

With both root pathogens, there were significant differences in tolerance exhibited by the lines. JI 181, JI 281, JI 2201 and JI 2551 showed good tolerance to both *Fusarium solani* and *Aphanomyces euteiches*. JI 188, JI 399, and JI 813 were the most susceptible to infection by both pathogens.
JI 188, JI 281, JI 399 and JI 813 were the most susceptible lines to downy mildew. Lines JI 1194, JI 2959 and Cooper were resistant to powdery mildew.

A notable outcome of this work was the characterisation of a source of resistance to downy mildew, for which genetic map information was obtained as part of a PhD study.

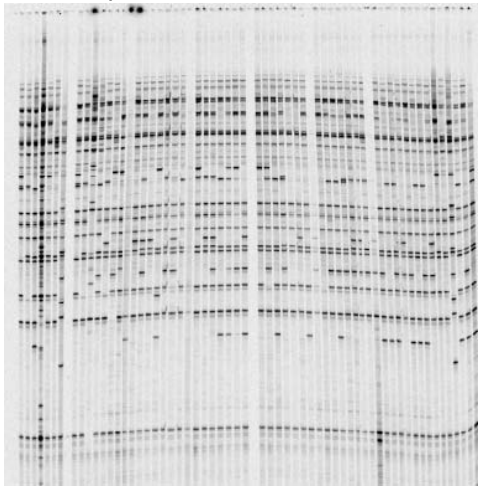
Objective 3: Performance

Extant varietal performance data will be associated with genotype data

The general aim of this objective was to generate a set of Recombinant Inbred Populations from a set of three cultivars chosen to have contrasting phenotypic traits and to be as different as possible genetically. We developed SSAP, a robust marker system for assessing pea genetic diversity (Ellis et al 1998, Vershinin et al 2003), and used this to assess genetic diversity among a set of 48 cultivars.

An example of one of these SSAP gel assays (PDR1 SSAP / Taq +AA) is illustrated in Figure 3.1 (right).

Figure 3.1 PDR1 SSAP analysis of PCGIN pea lines



Informative markers were scored on all of these lines, and their distribution among the lines was used to assess the pattern of genetic diversity by a variety of methods. The most direct method of displaying this is to undertake a Principal Component Analysis of the marker scores and compress this into a small number of dimensions that explain a high proportion of the total variance.

Eliminating the lines from the JIC *Pisum* germplasm collection meant that some of the informative markers were now monomorphic in this subset of 45 lines there were 153 informative markers. This is illustrated in Figure 3.2 for a subset of the 48 lines that excludes exotic material from JIC *Pisum* germplasm collection and therefore displays variation within cultivars:

These data were combined with an analysis of performance data from NIAB (Figure 3.3) trials in order to select an informative set of parental lines (see Annex 3.1 a poster at the 6th European Conference on Grain Legumes).

Figure 3.2 Principal Component Analysis of Marker scores for PCGIN pea lines

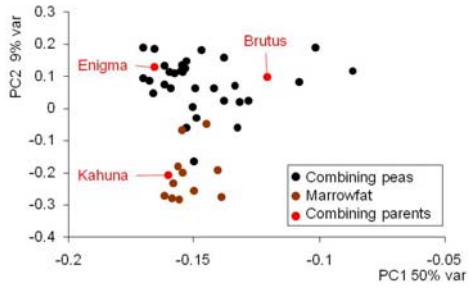
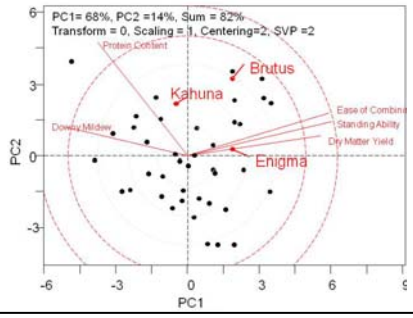


Figure 3.3 Biplot of phenotypic data for PCGIN pea lines



A compromise was required that conserved phenology while maximising other phenotypic differences and genetic diversity. As a result the lines Brutus, Kahuna and Enigma were selected. All are semi-leafless types and have recently been on the recommended list of cultivars.

Reciprocal crosses were made in all pairwise combinations and the progeny selfed as in single seed descent to generate three related F6 RIL populations with a total of 550 RILs available at F6.

A comparable study was undertaken with vining pea lines as an additional objective and RILs are being generated with Avola, Cabree and Waverex as parents. These will be advanced to RILs within the second phase of PCGIN.

The current status of the individual populations and sub populations is summarised in Table 3.1.

Table 3.1 Status of six PCGIN RIL populations

Cross (♀ x ♂)	**generation**	**Number of lines**
Brutus x Enigma	F6 (95)	10F5,1F3, 4D
Enigma x Brutus	F6 (86)	19F5, 5D
Brutus x Kahuna	F6(105)	2F5, 3D
Kahuna x Brutus	F6 (99)	7F5,3F4,1F3
Enigma x Kahuna	F6 (82)	13F5,10F4, 5D
Kahuna x Enigma	F6 (83)	20F5, 5F4, 2D
Waverex x Cabree	F3	110
Cabree x Waverex	F3	110
Waverex x Avola	F2	110
Avola x Waverex	F2	110
Avola x Cabree	F2	110
Cabree x Avola	F2	110

Objective 4: Reverse genetics

TILLING *for genes that regulate the development of the aerial part of the legume plant*

The PCGIN project was associated with a wider european project (Grain Legumes FOOD-CT-2003-506223) that developed many genetic and genomic resources for pea including a TILLING population for pea (Dalmais et al 2008, http://urgv.evry.inra.fr/UTILLdb) and a pea BAC library, developed at URGV Evry, with 12x genome coverage (see Hofer et al 2009, and GLIP 2008).

For PCGIN this provided us with DNA sequence information and TILLING mutants for the *Tendril-less* gene (Hofer et al 2009) and for our candidate genes for *Afila* (see Objective 5). Mutants are currently being screened for our candidate gene corresponding to *Apulvinic* locus. In Objective 6 mutants were generated for the pea Trypsin Inhibitor gene. These are described within the descriptions of Objectives 5 and 6 respectively.

Objective 5: Genetic mapping in crop legumes

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

The general aims of this part of the project were twofold: First we wanted to determine whether it would be practicable **to generate a comprehensive set of mutant** lines for forward and reverse genetic screens in pea. We know from other studies (Kaló et al 2004 MGG 272: 235-246, Choi et al 2004, PNAS101: 15289–15294) that the gene inventory of pea is likely to be very similar to that of the sequenced genomes of *Medicago truncatula* and *Lotus japonicus* so about 35,000 gene knock-outs are required for a comprehensive mutant population. In turn we knew from the analysis of two Fast Neutron (FN) mutants we had shown to be allelic to genes we had characterised previously (*unifoliata,* Hofer et al 1997 and *crispa,* Tattersall et al 2005) that we could not find any trace of the gene in PCR screens of the mutant lines (see Sainsbury, et al 2006) consistent with a deletion of the same scale as the size of a gene. This suggested that AFLP (Vos et al 1995) screens in pools of mutants should be a way of finding DNA sequences missing in FN lines but present in wild type progenitor stocks. At the time of submitting the PCGIN proposal we knew that this method had succeeded in isolating a candidate gene for *Tendril-less.*

From this we formulated several questions to ask of the FN population.

1) Can we estimate how many deletions there are in the population? (i) As a whole, (ii) per line and (iii) whether lines that do not have an obvious mutant phenotype are simply wild type 'escapes' from the FN mutagenesis.
2) Is there a better screen than AFLP on genomic DNA? Specifically we were concerned that if a FN deletion did not include the rare cutter site of the AFLP that we would miss some mutants. We are interested in screens that affect genes, so we investigated whether cDNA-AFLP could be a better approach. In this method a rare cutter site is added to the 3' end of a transcript so all expressed genes can be assayed.
3) Can we demonstrate that the candidate for *Tendril-less* corresponds to this gene. That is can we be sure that this is not simply an adjacent gene removed in a large deletion?
4) Can we use this method to isolate additional candidate genes?

1) The frequency distribution of deletions per line.

We undertook AFLP screens of a subset of lines including those showing obvious mutant phenotypes and those not easily distinguished from wild type. Eight of each class were screened with all combinations of primers for one Pst primer, corresponding to the observation of 2638 markers.

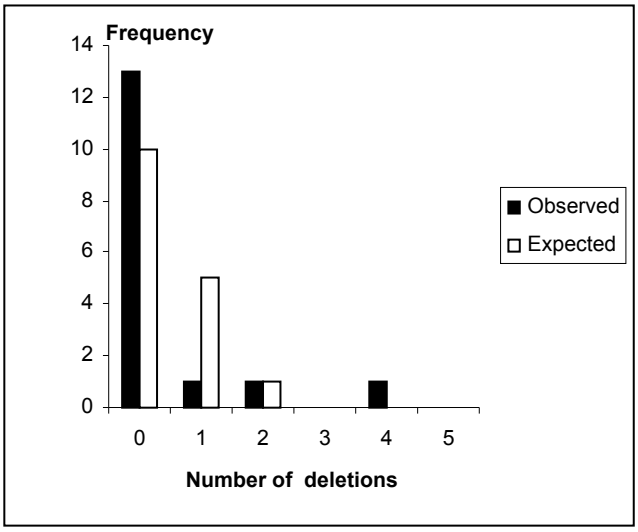


Figure 5.1 The frequency distribution of the number of missing amplicons in FN deletion pea lines.

The expected number is calculated from a Poisson distribution given the number of observed missing amplicons. A χ^2 test by category was not significant.

This analysis showed that the number of missing amplicons per line was highly variable, but that both classes of lines carried deletions. The number of deletions per line was low suggesting that artefactual candidates would be found rarely and that false leads would be rare enough to be resolvable by co-segregation tests. We could also estimate the total number of deletions in the population as a whole was about 20,000 implying about 0.6x genome coverage. The generation of a comprehensive ordered deletion set for pea is therefore feasible.

2. Is there a better screening method?
We compared the results of AFLP and cDNA-AFLP with a known deletion (at *Tendril-less*). While this showed that the cDNA-AFLP approach could work it was very cumbersome requiring a complex preamplification procedure and was also likely subject to bias in part from transcript abundance and the choice of frequent cutting site. Our conclusion from this was that genomic AFLP was the method of choice.

3. Demonstrate that we had identified the *Tendril-less* gene.
Four approaches were taken to confirm the identity of the Tendril-less gene as a novel Class I HD-ZIP transcription factor:
(i) We sequenced this gene in other *tl* alleles.
(ii) We obtained TILLING mutants of the gene and confirmed they had the expected mutant phenotype.
(iii) We showed that this gene was associated with tendril formation in other species.
(iv) We showed that the transcription of this gene was regulated as expected and found in the expected location.
These results are presented in Hofer et al (2009).

4. Isolation of additional candidate genes.
The success of this method for *Tendril-less* coupled with the rarity of deletion events per line suggested we could screen for further genes, and accordingly selected the lines with mutations corresponding to *Arthritic* [*Art,* 4 alleles] (has swollen nodes that may correspond to a gene involved in stem strength), *Apulvinic* [*Apu,* 4 alleles plus one found later] (which has altered leaflet attachment and may affect leaf movement and in turn modulate water use and light perception) and *Stipules reduced* [*St* one allele].

Art AFLP screens found multiple differences between independent *art* alleles and wild type or other mutant alleles. We remain uncertain of whether these differences are due to DNA methylation changes but our current working hypothesis is that this gene is included within a large deletion. Because of the limitation in time this was not pursued as we considered it likely to be difficult to identify a candidate gene easily.

St For *st* a clear AFLP band difference was found and this corresponded to *Rubisco activase*. Although this gene maps in the correct position it seems an unlikely candidate for *St* and no sequence differences were found between wild type and pre-existing *st* alleles. In conclusion *St* likely corresponds to an adjacent gene.

Apu Pursuing likely transcription factor candidates from synteny studies with *Medicago truncatula* led us to a complex family where we could not find a consistent difference between *St* and *st* alleles. However a linked paralogue was clearly deleted in four *apu* FN mutant lines. These map to the expected position and significant sequence differences between pre-existing *apu* mutants and their progenitor lines were found confirming this as a strong candidate for *Apu.* However a fifth FN line with the apulvinic phenotype appeared to contain this gene in an unaltered form. We are currently checking that this is indeed allelic to *Apu*. In addition primers specific to the *Apu* candidate gene have been designed and have been submitted to the TILLING resource at URGV Evry. Once these results are clarified we will describe this gene and its relationship to *Apu*.

B Biochemical studies and examination of the *Medicago truncatula* genome sequence suggested a candidate gene for *B*. This pigmentation mutant is not, of itself, a candidate gene of major agronomic importance, but in the FN population we identified an unstable allele of *B* (Figure 5.2). This seemed a likely source of a transposon that could be used in gene tagging experiments for pea, so we undertook the characterisation of this gene. We confirmed that the gene is missing in independent (stable) b FN alleles and exhibits significant structural differences (including a partial gene deletion) in pre-existing *b* alleles as compared to their progenitor wild type lines. We are currently preparing this for publication and will present the identity and sequence of these alleles at that time. The genetic instability has proved difficult to define but may correspond either to a recurring deletion or an insertion of a large DNA element in an intron, the (frequent) excision of which generates genic deletions. These studies are under way.

Figure 5.2. An unstable allele at the *B* locus

5B Integrated Genetic Maps
We intended to do two things regarding genetic maps. One was to confirm and extend gene specific mapping in pea that connects the pea genetic map to the *Medicago truncatula* genome sequence. This has two parts: Mapping genes and mapping markers for the end points of the *Medicago truncatula* euchromatin sequence (thus identifying regions of the pea genome where there is not a corresponding contiguous *Medicago truncatula* genome sequence. This is done in collaboration with Rene Guerts at the University of Wageningen and uses a wide cross RIL population (between *P. sativum* and *P. abyssinicum* that is not yet published).

Euchromatin / heterochromatin boundaries
We have identified markers corresponding to the following *Medicago truncatula* genomic locations:

Table 5.1 Pea map positions of end points of Medicago euchromatin sequence

Medicago truncatula	Pea location	Expected [Y/N]
Mt 5 South telomere	Lg I top	Y
Mt 5 South pericentric heterochromatin border	Lg I central	Y
Mt 1 North telomere	Lg II central	N
Mt 1 North pericentric heterochromatin border	Lg II central	Y
Mt 6 South pericentric heterochromatin border	Lg III upper	N
Mt 3 South pericentric heterochromatin border	Lg III middle	Y
Mt 2 North pericentric heterochromatin border	Lg III lower	Y
Mt 5 North telomere	Lg IV top	N
Mt 7 South telomere	Lg V top	Y
Mt 7 South pericentric heterochromatin border	Lg V middle	Y
Mt 4 South pericentric heterochromatin border	Lg VII internal	Y
Mt 4 North telomere	Lg VII internal	N (1)
Mt 4 South telomere	Lg VII bottom	(2)

Notes:
(1) The presence of a Nucleolus Organiser Region at this approximate location is a major structural difference between the pea and *Medicago* genomes and may account for this internal location.
(2) This marker does not map well and may correspond to a duplication associated with the satellite distal to the NOR in pea.

In general these locations show inconsistencies not expected given the general colinearity between these two genomes. However, whole arm rearrangements are common structural variants and these locations mark positions expected to be rich in repetitive elements in both genomes, so these inconsistencies may mark the borders of synteny blocks rather than identifying individual points of correspondence. It is notable that the Mt 1 North pericentric heterochromatin border and Mt 1 North telomere markers map to adjacent positions on the pea map, suggestive of such associations.

Comparative gene mapping
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, GLIP with which PCGIN has been strongly associated. (This information is publically available at: http://bioweb.abc.hu/mt/pisprim and http://cgi-www.daimi.au.dk/cgi-chili/GeneticMarkers/table). In collaboration with JIC, ~300 of these have been mapped in a JIC mapping population. Within PCGIN further genetic mapping with these markers with the aim of refining the position for a selection of 40 markers for genetic studies has been undertaken and the corresponding map is shown in Figure 5.3 below. These results confirmed the colinearity that is already published (Kaló et al 2004, Choi et al 2004).

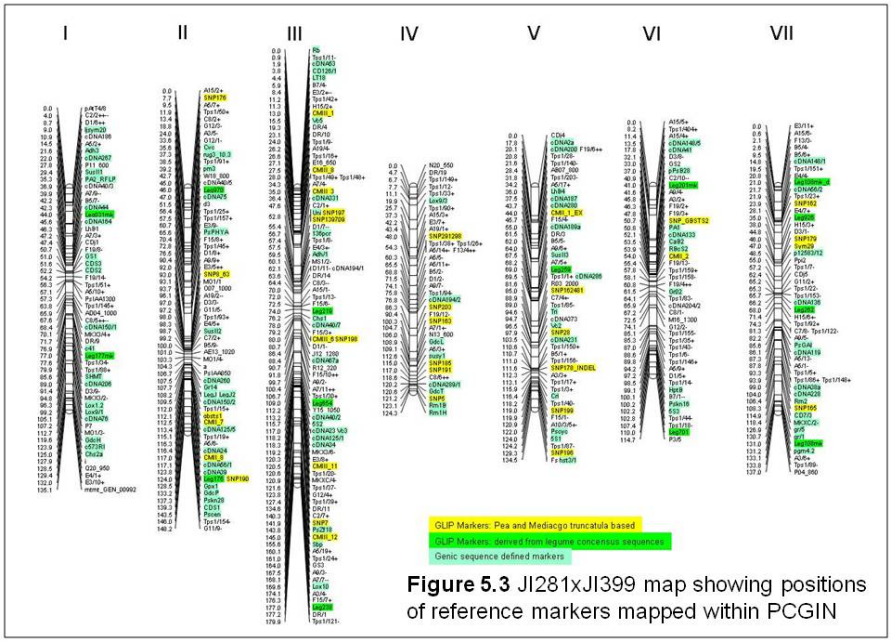
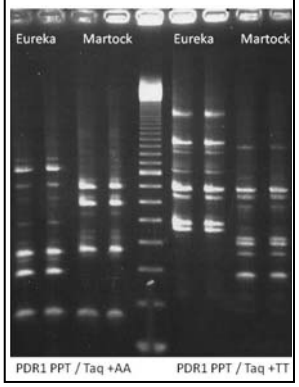


Figure 5.3 JI281xJI399 map showing positions of reference markers mapped within PCGIN

Faba bean

A secondary objective for PCGIN was to initiate approaches to genetic mapping in Faba bean. To this end a set of 95 faba bean (plus one *Vicia narbonensis*) accessions was assessed to determine the pattern of inter- and intra-varietal polymorphism with a set of 33 gene specific primers. Nine primer pairs gave an amplification length polymorphism within *Vicia faba*, 2 gave an ALP between *Vicia faba* and *Vicia narbonensis* and 2 SNP markers were found. Seven primer pairs showing ALP were scored on the full set of 96 lines. It appeared that the amplification obtained for the 2 individuals from the same line was different at a high frequency and for all the markers. This difference was reproducible suggesting a high degree of heterogeneity within these lines. The proportion of lines that were heterogeneous for a given marker was 0.10±0.02, but among 192 individuals scored at 3 loci only 5 heterozygotes were detected (and all at the same locus). For these loci the average inbreeding coefficient is 0.98±0.02, consistent with occasional outcrossing followed by selfing.

Figure 5.4 S-PCR based *PDR1* SSAP comparison of two *Vicia faba* genotypes



The genetic heterogeneity of faba bean cultivars presents problems for the generation of genetic stocks. Despite this there is abundant polymorphism in faba bean for example as evidenced by a modified SSAP approach that uses S-PCR (modified from Spertini et al 1999) and ethidium bromide staining (Fig 5.4, right). Duplicate samples from the same individual from cv. Eureka and Martock are compared for two primer combinations that reveal the technical reproducibility of *PDR1* SSAPs in *Vicia faba* and show the expected abundance of insertion site polymorphism.

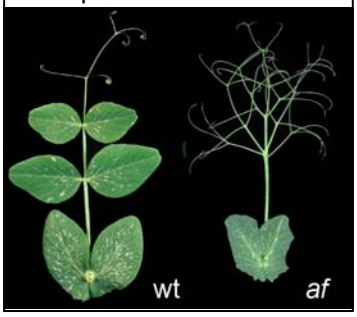
SSAP method

The *PDR1* SSAP method has been widely used for genetic mapping and diversity analysis in pea.
The method has used sequencing gels and ^{33}P labelling of a *PDR1* PPT-specific primer. As discussed above, this method can be replaced by a non-radioactive method using agarose gels where specificity is obtained using an S-PCR approach. This method is simple and robust, but agarose gels do not have high resolution for DNA molecules in the size range of these expected amplicons. Accordingly a method that uses fluorescent primers and resolves amplicons by capillary electrophoresis has been developed and is in press subject to minor editing. We have shown that this approach has the potential to be used as a co-dominant marker system at least for some loci.

Afila

The *afila* mutant (whole leaf compared to wild type, Fig 5.5, right) has been widely incorporated into pea breeding programmes and now dominates the combining pea market and is being introduced also into marrowfat and some vining varieties. The identification of this gene has been a target at JIC for some time and its identification will allow us to determine the extent of linkage disequilibrium in this region of linkage group I associated with the recent introduction of *af* alleles from a restricted germplasm. The approach we took to the identification of *Af* was comparative fine mapping.

Figure 5.5 Wild type and *afila* pea leaves



The *Af* locus maps to the bottom of pea linkage group I at a position corresponding to the top of *Medicago truncatula* chromosome 5. In the JI1201 x JI813 RIL population we identified several F6 lines that segregated for *Af* vs. *af.* F5 seed was collected from these lineages and selfed, obtaining 90 F6 seed of which 51 were heterozygous and these were selfed to give 967 F7 RILs used for fine mapping. The pea gene *GdcH* maps very close to *Af* and was used to identify *Medicago truncatula* BAC clones which enabled us to identify a contig from BAC fingerprint data. A further eight BACs were sequenced as part of the GLIP contribution to the Medicago genome sequencing project so a fully sequenced segment of the *Medicago truncatula* genome in the region of *Af* was identified. Primer pairs were designed from predicted ORFs within this region and used for further mapping.

Of 34 markers tested, 22 generated amplicons from pea and 10 of these were polymorphic in the JI1201xJI813 mapping population. Two genes were identified that flanked *Af* on either side and had a single recombination event within this expanded RIL population delimiting it to a 0.1cM$_{(Haldane)}$ interval. These two genes, *GdcH* (CR931738_19) and an HDZIP transcription factor (CT573020_12) identify a region of 390kb in *Medicago truncatula* that contains 60 annotated genes. Candidates were examined for differences between corresponding *af* and *Af* lines until a duplicate pair of transcription factor genes were found that were present in wild type but absent from *af.* This was confirmed in an analysis of 115 *af* lines from the JIC *Pisum* germplasm collection. Three lines were found that lacked only one of these genes (JI3504, JI3509 & JI3129) but these have a slightly different phenotype with occasional leaflet pairs on the rachis that also carries branched systems of tendrils.
(See Annex 5.1 for a poster on this subject presented at the 6th European Conference on Grain Legumes)

These results encouraged us to obtain TILLING mutant alleles of each gene. These were obtained and for one gene a STOP mutant was found and for the other two mis-sense mutations that would be predicted to have a major deleterious effect on protein function. None of the (homozygous) TILLING mutants had an obvious afila

phenotype. This suggests that both genes need to be defective for the mutant phenotype to be revealed. In a collaboration with INRA Versailles we are attempting to obtain such plants with Virus Induced Gene Silencing (Constantin et al 2004).

Objective 6: Genetics of Seed Quality Traits

This objective has linked closely with Objective 1, having extensive interaction with the wider stakeholder group that includes end-users as well as breeders. Improvements to home-grown sources of animal feedstuffs are relevant to Defra's objectives for improved quality, linked to a reduction in processing, energy use and minimisation of waste. Improved quality in feed, coupled with improved digestibility, will be linked to a reduction of potentially eutrophicating waste, with associated effects on natural resources. An improved scientific understanding of food quality is linked to Defra's overall objectives for improved food quality and nutritional value, coupled to a reduction in processing and waste. Improved quality selection procedures will be linked to increased efficiency, and reduced energy use, waste and carbon emissions. Biological solutions to problems with quality can avoid the use of undesirable chemicals that may not be available for use in the future.

Following extensive consultation and discussion at early management and stakeholder meetings, traits relevant to end-use were defined, with the aim of dissecting their genetics. The definition and prioritisation of target traits, coupled with an understanding of their genetics, had the broad objective to provide tools and resources to industry. Importantly, the aim was to develop and use simplified genetic screens, which are particularly relevant to the screening of seed characteristics at a very early plant stage and, where appropriate, to identify novel variants that could be exploited in breeding programmes.

A: Animal feed

The so-called anti-nutritional seed proteins in pea, the trypsin-chymotrypsin inhibitors (TI/CTI) encoded by genes at the *Tri* locus and pea albumin 2 (PA2), were chosen as proteins that impact directly or indirectly on end-use and whose reduction or absence could improve quality. At the stakeholders' meeting in October 2007, information provided by Mick Hazzledine, Premier Nutrition, was presented to the meeting, to provide a current industrial perspective on crops for animal feed.

The TILLING platform (http://urgv.evry.inra.fr/UTILLdb) has been exploited to identify mutants in the two seed protein gene families.

For the TI proteins, the mutants included three mis-sense mutations giving rise to amino acid changes:

C77Y (TI 2808): a highly conserved cysteine residue has been replaced by a tyrosine.

S85F (TI 671): the active site serine in the chymotrypsin inhibitory loop has been replaced by a phenylalanine.

E109K (TI 895): a glutamic acid residue in the carboxy-terminal extension has been replaced by a lysine, an alteration in overall charge in this region of the protein that is subject to post-translational processing.

For the PA2 proteins, three unique mis-sense mutations were identified:

G138E (PA2 2628): a glycine residue has been replaced by a glutamic acid.

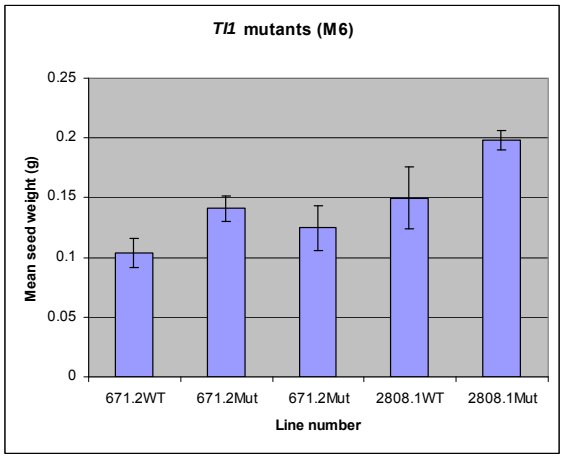
T175M (PA2 1174): a threonine has been replaced by a methionine.

E186K (PA2 487): a glutamic acid residue has been replaced by a lysine. The first and third of these mutations involve an alteration in overall charge.

Although mutant lines were identified readily at M3 and M4 generations, the effects of the mutations could not be assessed prior to backcrossing the mutant lines. Negative effects on seed size and seed yield, not associated with the mutation of interest, were apparent for several lines (mean seed weights of 30-60mg for some PA2 lines, compared with 200mg for the parent cv. Cameor), implying that the background mutation load for the population is relatively high. The differences in seed size for the inhibitor gene (*TI1*) mutants compared with segregants having wild type alleles are shown in Figure 6.1.

Assays performed on the M5 (non-backcrossed) seeds indicated a significant effect on the activity of the TI 2808 family but, due to the effects of additional mutations, the control and mutant seeds cannot be compared strictly. All the mutant lines were therefore backcrossed to the parent, cv. Cameor, mutant lines selected and the process repeated to give BC2 mutant lines. Bulked mutant seeds from the second backcross generation are now available for all the lines, apart from TI 671 and PA2 1174, which are at BC2F1 and BC2F2 segregating seed stages, respectively. Analysis of the PA2 mutant lines will complement the metabolomic analyses that have been conducted using a naturally occurring mutant (Vigeolas et al., 2008).

Figure 6.1: Differences in seed size for two TI mutant (Mut) segregant lines, 671 and 2808, compared with segregant lines having the wild type alleles (WT). Values are based on the mean of three plants for every line, and a yield of 5-11 seeds per plant**.**



B: Food crops

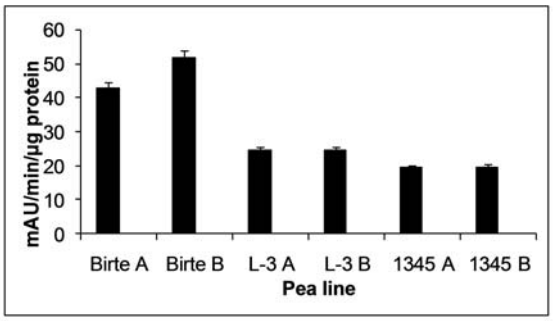
The genetics of green colour retention by seeds

The retention of green seed colour by vining and marrowfat peas is an important quality parameter. If colour is lost, seeds can be downgraded, leading to income loss as well as undesirabler artificial colour enhancement of food products. Genetic loci implicated in colour retention have been targeted, therefore, to identify variation in gene structure, and to determine what the consequences of such variation are for gene expression and phenotype.

Impact of variation in seed lipoxygenase activity on seed colour retention:

Lipoxygenase-3 (Lox-3) has been shown to be capable of bleaching pigments through co-oxidisation by products of polyunsaturated fatty acid oxidation (Porta and Rocha-Sosa, 2002). Two variant (yellow-seeded) lines with reduced Lox-3 activity were identified (L-3 and H53). H53 is a RIL selected from a cross between a cultivar and a wild pea, JI 1345, where the latter and H53 both showed reduced amounts of both Lox-3 and Lox-2 proteins. (Lox-2 is implicated in the generation of off-flavours and hence organoleptic properties of foods). L-3 and JI 1345 showed 52% and 41% of the Lox enzyme activity of cv. Birte (Fig. 6.2). Due to the reduction in two Lox proteins (Lox-3 and Lox-2) that are the products of linked loci, lines may be generated that have improved organoleptic properties in addition to reduced seed colour bleaching.

Figure 6.2: Lox activity in seeds from the pea lines, Birte, L-3 and JI 1345. A, B = biological replicates; SD of replicate assays.



Crosses were established between L-3/H53 and a number of marrowfat lines representative of those used by the industry. Marrowfat lines are prone to undesirable seed bleaching, not evident in all green-seeded lines. Progeny of genotype *i af* (green seeded and semi-leafless characters from the marrowfat parent) were selected to screen for *Lox* gene variants. Primers based on the pea *Lox-3* gene promoter (EMBL accession number X78581) were used throughout to select the gene variants. The primers were designed on a promoter region that showed size variation among lines (Fig. 6.3). Among marrowfat lines, two size variants were observed, the two Lox variant lines differed from the marrowfats and from each other, and the two 'supergreen' (see below) lines showed further variation (Fig. 6.3). Agarose gel electrophoresis was sufficient to distinguish the three marrowfat having the small amplicon, whereas acrylamide gels were required to distinguish all other variants.

F_3 plants of genotype *i*, *af*, Lox^{var} were used for backcrossing to their respective marrowfat parent and BC1 F_2 plants carrying the variant *Lox* allele were selected (see for example Figure 6.4). A small plot trial was conducted in 2008, using BC1 F_3 lines (Lox^{L-3}) from the Kahuna and Princess crosses. Three replicates of the BC1 lines were compared with the cultivar parent for seed bleaching under field conditions. Bleaching pressure was considered to quite high towards the seed maturation stage, i.e sunny, wet conditions, and bleaching was observed for seeds from all replicates. However, the BC lines having the L-3 allele appeared to be visually more green than the corresponding marrowfat parent. A SPECTROmatch (Sheen Instruments) was used to quantify the colour differences between samples.

Figure 6.3: Variation in *Lox-3* gene fragments amplified from 10 pea lines using a pair of promoter primers and visualised on an acrylamide gel. Marrowfat lines and genotypes from the JI germplasm collection (JI) are indicated in brackets. LL indicates low Lox (L-3, JI 1345) and SG 'supergreen' lines. - = no DNA control. DNA markers are shown in the first track (left).

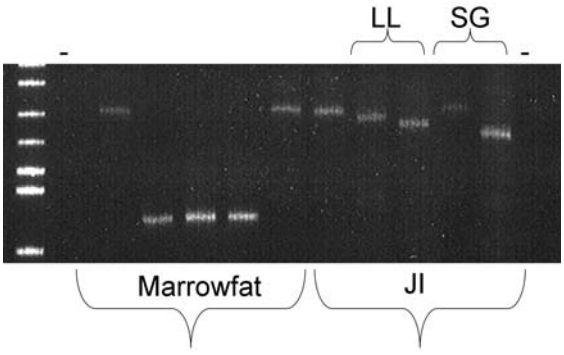


Figure 6.4: Screening Samson x H53 lines (bracketed) for *Lox* gene variants by PCR. The two tracks on the right show the two parents, Samson and H53, respectively (left to right). The F_2 lines having the H53 allele (two lines) are likely to be homozygotes, whereas the remaining F_2 lines will be either homozygous for the Samson allele or heterozygotes. DNA markers are shown in the first track (left).

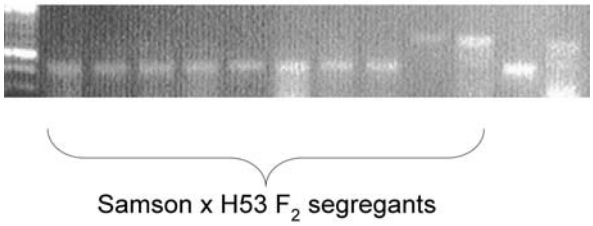


Figure 6.5: Reflectance curve from 400-700nm determined for ground meal for different groups of peas. Y = yellow-seeded, BM = bleached marrowfats, UM = unbleached marrowfats, SG = supergreen lines.

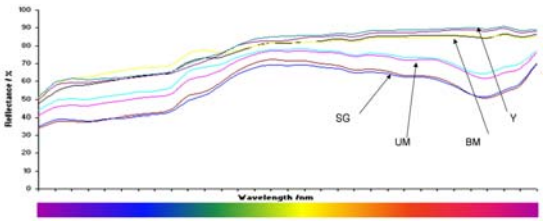
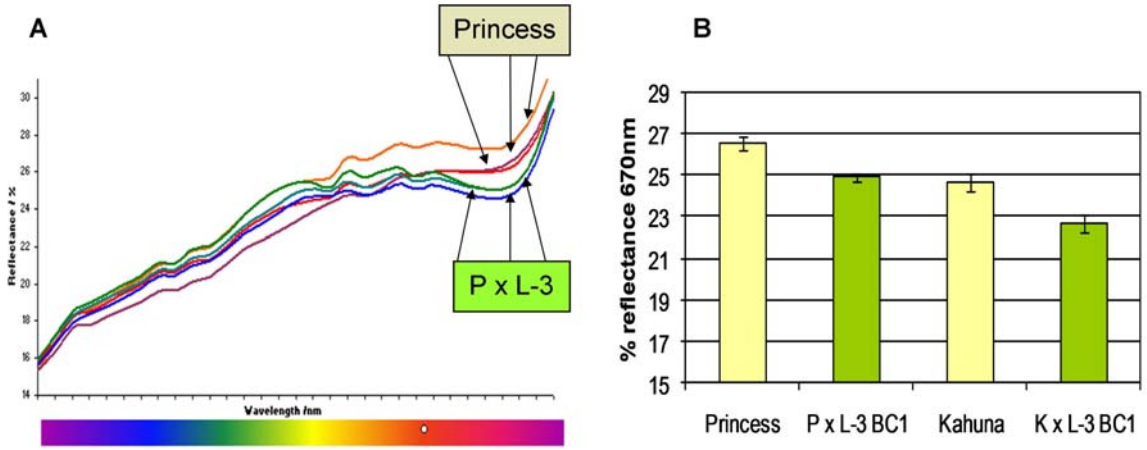


Figure 6.5 shows that the spectra obtained for meal obtained from genetically yellow seeded and bleached green seeds were very similar. Unbleached green seeds formed an intermediate group, whereas 'supergreen' lines showed the lowest values. The same assay was applied, therefore, to the whole seed samples obtained from the field, where there was much less uniformity of sample and seed testa colour also contributed to the spectrum.

Figure 6.6: Reflectance curve from 400-700nm determined for whole seeds obtained for 3 replicates of cv. Princess and the corresponding backcross line having the variant Lox allele (**A**). The % reflectance at 670nm for the two pairs of marrowfat lines are plotted in (**B**).



The data from this experiment (Fig. 6.6) suggest that there is a small but significant difference ($p < 0.05$) between the members of each pair, and a clear distinction of triplicates at 670 nm in Fig. 6.6A. The data suggest that, although bleaching was not abolished, it was reduced in lines with lower Lox activity. Clearly these experiments need to be repeated following a further backcross. The BC2 seeds are now available for the cv. Kahuna cross and will be trialled in 2010, alongside a second equivalent pair, involving cvs. Princess or Samson as parents.

The basis for green colour stability has been investigated furthermore in 'supergreen' lines that are bleach resistant (discussed below). *Lox* gene variation is also apparent in two 'supergreen' lines (Figure 6.3), providing

additional genetic variation that has been investigated with respect to enzyme activity. Although one of the 'supergreen' lines has shown a reduced seed Lox activity (approximately 90% of that of wild type, data not shown), the second (JI 2621) did not, suggesting again that this enzyme activity does not explain all the variation observed for colour stability. This is further supported by analysis of F3 families from a cross between marrowfat lines and a supergreen line, JI 2621. Here bleach resistance was clearly visible among F2/F3 seeds (Fig. 6.7).

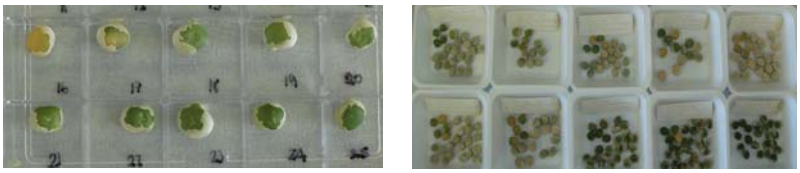


Figure 6.7: Variation in seed colour bleaching among F_2 seeds (left panel) and corresponding F_3 progeny (right panel) derived from the cross Samson x JI 2621.

Colour analysis, performed on F4 seed meal from F3 families classified according to *Lox* allele and phenotype (bleached/unbleached), showed that colour retention was higher among those families having the JI 2621 *Lox* allele (Figure 6.8). These data again support a partial effect of Lox on colour retention. However, the total Lox activity was not lower in JI 2621 than in cvs Kahuna or Princess (see above), indicating that a genetically linked locus rather than *Lox* itself may explain the effect on colour retention. An alternative explanation is that Lox-3 activity is lower in JI 2621 but that the total Lox activity is not, due to a higher activity of Lox-2 or related enzyme. Clearly, the development of recombinant inbred lines (RILs) based on parents with contrasting colour retention, and the QTL approach outlined below, will clarify the contribution made by individual loci to the bleaching phenotype.

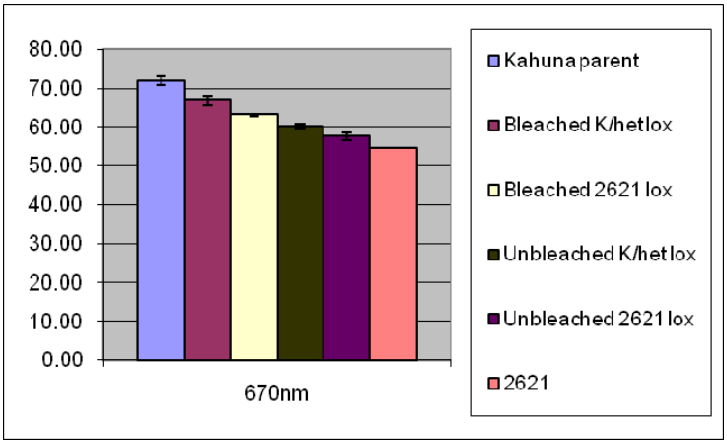


Figure 6.8: The % reflectance at 670nm of seed meals of progeny from F3 families derived from Kahuna x JI 2621 and classified according to *Lox* allele and phenotype. The supergreen and marrowfat parent values are the extremes (right, pink and left, blue, respectively). Among the four progeny classes, those carrying the Lox^{2621} allele showed higher colour retention than the corresponding lines that had the Lox^{Kahuna} allele. Means and SE of three replicate families.

Impact of a 'supergreen' phenotype on seed colour retention:

Two pea lines having an intense green colour, which appears to be stable in stored seed samples, were selected from the JI germplasm collection. These were crossed to one or more of the marrowfat lines, Princess, Kahuna and Samson, which are prone to bleaching. (Kahuna has been retained as a common parent here and for Lox crosses, as this cultivar is also one of the parents in the three-way cross of combining peas in Objective 3). One of the crosses was selected for the development of RILs (Princess x JI 185, see below).

Allelic variation at the cotyledon colour locus, *i*, was examined within germplasm accessions and marrowfat lines, using primers based on the *sgr* (staygreen) gene (Aubry et al., 2008). Considerable variation for this gene exists within green-seeded lines. None of the marrowfat lines examined showed any amplification using these primers. Figure 6.9 shows a selection of green and yellow seeded lines examined. Some of the other green- and all yellow-seeded lines examined yielded a fragment (Figure 6.9). This screen provides useful markers to follow the inheritance of *i* gene variants in some crosses and to reveal whether or not this variation impacts on the activity of the enzyme(s) involved in breakdown of colour pigments. One of the 'supergreen' lines (JI 185) yielded a fragment, whereas the other did not (Figure 6.9), facilitating the screening of crosses involving JI 185.

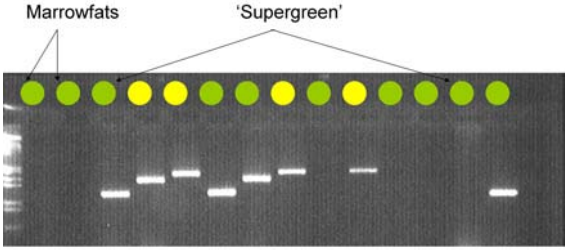


Figure 6.9: Amplification by PCR of *sgr* gene fragments in a selection of pea lines, including marrowfat (arrowed), vining, 'supergreen' (arrowed) and yellow-seeded lines. Colour codes indicate if lines are green or yellow-seeded. DNA markers are shown in the first track (left).

Homozygous families having the JI 185 *sgr* allele have been selected from the Princess x JI 185 and Kahuna x JI 185 crosses, based on screening F_3 plants by PCR. Two backcrosses have been carried out to the marrowfat parent to provide material suitable for field evaluation of the effect of introgression of the JI 185 *sgr* allele.
Sequence analysis has shown no consistent difference in the *sgr* gene between yellow- and green-seeded lines. Differences primarily associated with the large intron account for the grouping of some lines (Figure 6.10). Further analysis has included sequencing of the 5' part of the gene (dotted brown line in Figure 6.10) and approximately 600 bp of promoter. None of the available sequence discriminates seed groups by phenotype (green, yellow, bleach susceptible, 'supergreen').

Expression analysis has shown that sgr is expressed in all lines thus far examined, including representatives of all the phenotypic groups above and marrowfat lines. Therefore a large insertion in intron 3 may prevent the amplification of this part of the *sgr* gene in some lines, but not its expression. It is possible that quantitative differences in expression relate to phenotype, at least for the cotyledon yellow/green trait. Alternatively, genes close to *sgr* may be implicated in the colour degradation pathway and this will be explored further.

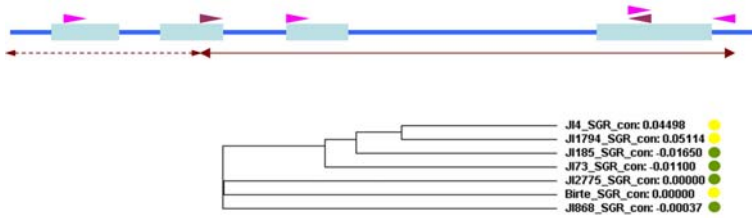


Figure 6.10: A schematic of the *sgr* gene, showing the exon 3/4 genic region that has been amplified (solid brown line). Exons are shown as blocks (pale blue) with introns as lines (dark blue). Primer positions are indicated as forward or reverse block arrows. The relationship between genotypic sequence variation and seed colour phenotype is shown as a tree.

Development of mapping populations for seed quality:
The Princess x JI 185 cross was chosen for the development of RILs for QTL mapping studies. Although JI 185 has a tall phenotype, the cross proved to be manageable, in contrast to those involving JI 2621 (extremely tall, much vegetative growth and high sensitivity to daylength delaying flowering).

One hundred and fifty Princess x JI 185 RILs at F6 were sown and leaf DNA samples used for Sequence Specific Amplified Polymorphism (SSAP) marker analysis, based on the *Ty1-copia* like element *PDR1* (Ellis et al., 1998, Jing et al., 2005), and using a fluorescent non radioactive marker assay (Knox et al., in preparation). Analysis of the two parents revealed a total of 191 potential SSAP markers, of which 130 have known map positions (Annex 6.1). Of the latter, 40 have been mapped in the RILs, along with the morphological markers, *af*, *le* and *a* and the functional gene markers, *sgr* and *cvcA* (EMBL accession number X06398), providing a genetic map that will be expanded by the addition of further functional and SSAP markers and that is available for QTL studies of seed quality.

F7 seeds from the RILs will be developed further as single seed descent lines, and bulked F8 seeds used for QTL analyses in the next phase of PCGIN. Field data will be collected in 2009 for the RILs (F7 bulks) grown under field conditions, in collaboration with the University of Saskatoon, Canada. Collaborations are in place with Limagrain UK to exploit the tools and resources that are being developed (see Objective 1).

Related projects on seed quality:
The identification of seed metabolites involved in quality, and how genetic markers associated with these may facilitate breeding for the food industry, was discussed at PCGIN meetings, led by Adrian Charlton, CSL, York. NMR analysis and metabolite profiling of mature seeds of a vining pea cultivar (*cv.* Puget) have been carried out within the PCGIN satellite project at CSL. The development of robust protocols for profiling seeds has allowed metabolites that respond to the environment as well as to genotype to be identified. A first paper based on the NMR profiling of pea leaves, that preceded the seed profiling analysis, has been published (Charlton et al., 2009). This work will be pursued within the LINK project, QDiPS (see Objective 1).

A student project (M.Sc. in Plant Genetics and Crop Improvement, JIC-UEA, Norwich) explored natural variation in defensin genes in pea, including an Australian bruchid-resistant *P. fulvum* line, in response to the growing problem with bruchid beetles observed for *Vicia faba* in UK and pea elsewhere in the world. These data are presented in Annex 6.2.

References to published material

9. This section should be used to record links (hypertext links where possible) or references to other published material generated by, or relating to this project.

Aubry S., Mani J. and Hörtensteiner S (2008) Stay-green protein, defective in Mendel's green cotyledon mutant, acts independent and upstream of pheophorbide *a* oxygenase in the chlorophyll catabolic pathway. Plant Molecular Biology **67**: 243-256

Charlton A.J., Donarski J.A., Harrison M., Jones S.A., Godward J., Oehlschlager S., Arques J.L., Ambrose M., Chinoy C., Mullineaux P.M. and Domoney C. (2009) Responses of the pea (Pisum sativum L.) leaf metabolome to drought stress assessed by nuclear magnetic resonance spectroscopy. Metabolomics **4**: 312-327

Choi H.-K., Kim D.-J., Zhu H., Mun J.-H., Baek J-M, Roe B., Ellis N., Young N.D., Doyle J., Kiss G. and Cook D.R. (2004). Conserved gene order between crop and model legume species.Proc. Natl. Acad. Sci. USA 101: 15289–15294

Constantin GD, Krath BN, MacFarlane SA, Nicolaisen M, Johansen IE, Lund OS (2004) Virus-induced gene silencing as a tool for functional genomics in a legume species. Plant Journal. 40:622-631

Dalmais, M., Schmidt, J., Le Signor, C., Moussy, F., Burstin, J.,Savois, V., Aubert, G., Brunaud, V., de Oliveira, Y., Guichard, C.,Thompson, R., and Bendahmane, A. (2008). UTILLdb, a *Pisum sativum* in silico forward and reverse genetics tool. Genome Biol. 9: R43

Duc G. (1997) Faba bean (*Vicia faba* L.). Field Crops Research **53**: 99-109.

Egan J. and Crouch J. (2009) Faba bean growing guide 2009.
http://www.sardi.sa.gov.au/__data/assets/pdf_file/0004/45958/080698_bean_press.pdf

Ellis T.H.N., Poyser S.J., Knox M.R., Vershinin A.V. and Ambrose M.J. (1998) *Ty1-copia* class retrotransposon insertion site polymorphism for linkage and diversity analysis in pea. Mol. Gen. Genet. 260: 9-19.

GLIP (2008) http://www.eugrainlegumes.org/documents/documents/Publishable_Final_report_v15.pdf

Hardie D.C., Baker G.J. and Marshall D.R. (1995) Field screening of *Pisum* accessions to evaluate their susceptibility to the pea weevil (Coleoptera: Bruchidae). Euphytica **84**:155–161

Hedley, C. L. and Ambrose, M. J., (1981) Designing 'leafless' plants for improving the dried pea crop, *Adv. Agron.*, **34**: 225-277.

Hofer J., Turner L., Hellens R., Ambrose M., Matthews P., Michael A. and Ellis N. (1997) *Unifoliata* regulates leaf and flower morphogenesis in pea.Curr. Biol. 7: 581-587

Hofer J., Turner L., Moreau C., Ambrose M., Isaac P., Butcher S., Weller J., Dupin A., Dalmais M., LeSignor C., Bendahmane A. and Ellis N.(2009) *Tendril-less* regulates tendril formation in pea leaves. The Plant Cell (2009) 21:420-428

Holden D. (2009) Genetic studies of wide crosses between cultivated pea species, and the domestication of *Pisum abyssinicum.* Ph.D. Thesis University of East Anglia

Holland M.R., Grace J. and Hedley, C.L (1991) Momentum absorption by dried-pea crops. 1. Field measurements over and within varieties of different leaf structure. Agric. and Forestry Meteorology **54**: 67-79.

Jing R., Knox M.R., Lee J.M., Vershinin A.V., Ambrose M., Ellis T.H.N. and Flavell A.J. (2005) Insertional polymorphism and antiquity of *PDR1* retrotransposon insertions in *Pisum* species. Genetics **171**: 741–752

Kaló P., Seres A., Taylor S.A., Jakab J., Kevei Z., Kereszt A., Endre G., Ellis T.H.N. and Kiss G.B. (2004) Comparative mapping between *Medicago sativa* and *Pisum sativum*. Molecular Genetics and Genomics 272: 235-246

Lawes D.A., Bond, D.A. and Poulsen, M.H. (1983) Classification, origin, breeding methods and objectives. In; the Faba Bean (*Vicia faba* L.) Ed. P.D. Hebblethwaite, Butterworths, London.

Link W. (1990) Autofertility and rate of cross-fertilization: crucial characters for breeding synthetic varieties in faba beans (*Vicia faba* L.). Theoretical and Applied Genetics **79**: 713-717

Liu Y.J., Cheng C.S., Lai S.M., Hsu M.P., Chen C.S. and Lyu P.C. (2006) Solution structure of the plant defensin VrD1 from mung bean and its possible role in insecticidal activity against bruchids. Proteins **63**: 777–786

Porta, H. and Rocha-Sosa, M. (2002) Plant lipoxygenases. Physiological and molecular features. Plant Phys.**130:** 15-21

Sainsbury F., Tattersall A.D., Ambrose M.J., Turner L., Ellis T.H.N. and Hofer J.M.I. (2006) A *crispa* null mutant facilitates identification of a *crispa*-like pseudogene in pea. Functional Plant Biology 33: 757–763

Shiau Y.S., Horng S.B., Chen C.S., Huang P.T., Lin C., Hsueh Y.C. and Lou K.L. (2006) Structural analysis of the unique insecticidal activity of novel mungbean defensin VrD1 reveals possibility of homoplasy evolution between plant defensins and scorpion neurotoxins. Journal of Molecular Recognition **19:** 441-450

Siddique K.H.M, Loss S.P, Enneking D. (1996) Narbon bean (*Vicia narbonensis* L.): a promising grain legume for low rainfall areas of south-western Australia. Australian Journal of Experimental Agriculture **36**: 53–62.

Snoad, B., (1974) A preliminary assessment of 'leafless peas', Euphytica, **23**: 257-265.

Spertini D., Beliveau C., Bellemare G., Biotechniques (1999) Screening of transgenic plants by amplification of unknown genomic DNA flanking T-DNA. BioTechniques 27: 308-314

Tattersall A.D., Turner L., Knox M.R., Ambrose M.J., Ellis T.H.N. and Hofer J.M.I. (2005) The mutant *crispa* reveals multiple roles for *PHANTASTICA* in pea compound leaf development. The Plant Cell 17: 1046-1060

Vershinin A. V. Allnutt T. R. Knox M. R. Ambrose M. J. Ellis T. H. N. (2003) Transposable elements reveal the impact of introgression, rather than transposition, in Pisum diversity, evolution, and domestication. Molecular Biology and Evolution **20**: 2067-2075.

Vigeolas H., Chinoy C., Zuther E., Blessington B., Geigenberger P. and Domoney C. (2008) Combined metabolomic and genetic approaches reveal a link between the polyamine pathway and albumin 2 in developing pea seeds. Plant Physiology **146**: 74-82

Young, J.P.W., Johnston, A.W.B., Brewin, N.J. (1982) A search for peas (Pisum sativum) showing strain specificity for symbiotic Rhizobium leguminosarum. Heredity **48**: 197-201.

Young, J.P.W., Matthews, P. (1982) A distinct class of peas (*Pisum sativum*) from Afghanistan that show strain specificity for symbiotic Rhizobium. Heredity **48**: 203-210.

Vos P, Hogers R, Bleeker M, Reijans M, Lee Tvan de, Hornes M,Frijters A, Pot J, Peleman J Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucl. Acid. Res. 23:4407-4414