General Enquiries on the form should be made to:
Defra, Procurements and Commercial Function (Evidence Procurement Team)
E-mail: research.competitions@defra.gsi.gov.uk

Evidence 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 Evidence 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
  An Evidence Project Final Report must be completed for all projects.
- This form is in Word format and the boxes may be expanded, 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: IF0147
2. Project title: Defra Pulse Crop Genetic Improvement Network
3. Contractor organisation(s): John Innes Centre, Norwich Research Park, Colney, Norwich, NR4 7UH
4. Total Defra project costs (agreed fixed price): £ 1528009
5. Project: start date 01/02/2009
   end date 2014

6. It is Defra's intention to publish this form.

   Please confirm your agreement to do so ..... **YES**

   (a) When preparing Evidence Project Final Reports 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 Evidence Project Final Report 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

We would like to discuss with Defra that the report contains unpublished data, some of which needs to be considered as confidential until publication.

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 Pulse Crop Genetic Improvement Network (PCGIN) was established as a platform for the genetic improvement of legume crops in the UK. It provides an environment of exchange for breeders, producers, end users and the research base, through which scientific knowledge, resources and results are delivered to add value to pulse crops. The project makes genetics expertise, resources, and their analysis publically available for use in breeding, and in the definition and improvement of product quality for commercial and public use. PCGIN investigates and executes the translation of genomic and genetic research to crop improvement, drawing from the input of leading UK producers, and fulfilling Defra objectives relating to sustainable agriculture. Furthermore, it provides links with, and involvement in, legume crop research programmes within Europe and world-wide. It continues to foster engagement with combining, vining, canning and related (fresh picking) industries, and acts as the umbrella network for the genetic improvement of both **pea** and **faba bean** legume (pulse and vegetable) crops, within a UK context but also within the rest of Europe and world-wide.

The scientific objectives of the network fall into three research areas, with an additional objective related to communication.

Objective 1: trait biology; to understand the genetic control of key traits that impact on plant performance and seed quality in pea. The focus has been on agronomic traits in the field, namely those traits that are components of yield determinants, and seed quality traits that determine crop value. Plant populations were established for which field performance could be associated with genetic markers, which in turn would form the basis for selection in breeding programmes.

Experiments for yield traits were conducted on three populations over three years at two sites, where each population contained over 200 progeny lines. Phenotypic trait analysis of populations segregating for economically important traits under field conditions has been used to define 'quantitative trait loci', those regions of the genetic map which influence better or worse performance. Defining genetic markers that are closely linked or adjacent to the genetic regions determining trait variation allowed the development of tools which are inputting directly to breeding programmes. The outputs will have application to related pulse species, where markers may be transferred in consultation with stakeholders.

In a similar way, markers associated with seed quality traits of economic importance have been defined. The genetic regions having the largest effects on seed size and colour retention have been identified and the role of candidate genes investigated in some detail. The data illustrate in a very powerful way the value of identifying and investigating variants of genes in the wider gene pool, typically from within germplasm collections, where variants have been collected from very distinct environments worldwide. It is commonly the case that there is very limited variation in breeders' lines.

The combined yield and seed quality data indicate that there is some co-localisation of these genetic determinants, and will allow breeders to incorporate genetic markers into apt screens that can be performed on very young plants, saving on space and time. This means that careful use of marker data is necessary to optimise selection for both groups of traits. In other words, continuous selection of large-sized seeds and those showing the best retention of colour by phenotype alone could lead to repeated selection of lines with lower yield overall. Knowledge of markers and the contribution made by each of the genes present in parent lines will lead to a step change in crop improvement for global traits.

This part of the project provided tools by which seed quality traits can be improved for a number of end uses: seed colour stability via novel variants of genes, nutritional quality though a reduction in anti-nutrients, and improved sugar-starch balance, aimed at export, feed and food markets, respectively. Some of the novel alleles were identified through collaboration with international programmes or by adoption of high-throughput screens of wider plant collections.

Objective 2: genetic mapping; to underpin the genetic improvement of both pea and faba bean, by providing reference maps for marker analysis of the traits under investigation. Dedicated PCGIN mapping populations were built and populated with genetic markers that supported trait analysis in Objective 1, and furthermore provide resources for genetic mapping in faba bean, where few or none have been available previously. The challenge was to use lines that were based on commercially relevant (breeding) lines, which by their nature have a very narrow genetic base; in faba bean, it is clear that diversity overall is very much lower than in pea. Here we also exploited the wider pea germplasm to provide variation in seed colour stability. Marker analysis exploited all available information for genetic loci determined previously for JIC mapping populations. Despite the low level of genetic variation overall in the commercial pea parents, where variation could be detected for gene or repetitive sequence-based markers, this could be assigned to a genetic location by extrapolation to the JIC mapping populations where there was much greater diversity and density of markers.

The genetic maps that underpin trait analyses in four PCGIN pea populations plus the relevant genetic stocks are available as a resource to breeders and end-users. The field experiments and trait analyses were performed using stocks multiplied from F6 lines; all recombinant inbred lines have been advanced to F10, providing greater uniformity within lines for future study. Representative variant lines are feeding into a large international study being coordinated by the University of Saskatoon, Canada.

In faba bean, two diverse lines, Albus and BPL10, have been chosen to provide genetic marker information to build the first genetic map for this crop relevant to UK trait analysis, including resistance to stem nematodes. This trait is being introduced into elite faba bean breeding material in collaboration with plant breeders and industry (Technology Strategy Board (TSB) co-funded project).

Objective 3: genetic resources; to expand the resources that are available for both pea and faba bean crops. We have exploited collaborations with other European platforms, national and international germplasm resources and international collections of both pea and faba bean. We have sought and characterised novel mutations associated with quality and agronomic traits and introgressed these into appropriate backgrounds to prove their involvement in plant processes. The utility of a fast-neutron mutant pea resource has been extended and the population used as the foundation for an industry-led project. This and a TILLING platform have provided access to completely novel resources to underpin trait improvement. Examples include loss of a protein with low nutritional quality, reduction of the major anti-nutritional proteins by over 50%, downy mildew resistance, and potential enhanced seed colour stability in pea.

A collection of over 900 *Vicia faba* accessions, comprising a broad range of UK, EU and Middle Eastern cultivars and landraces, has been acquired for long-term seed storage at NIAB. More than 130 accessions have been entered into a single seed descent (SSD) programme to obtain inbred germplasm. This is work in progress in parallel with marker development to assess the progressive reduction in heterogeneity (or increasing homogeneity) within the programme. The collection provides a diverse range of homogenous germplasm suitable for further trait analysis and marker development. A highly inbred line of Hedin has been proposed as a reference faba bean genotype for PCGIN and UK research. This line (> 30 generations of selfing) was obtained from the University of Gottingen and is being multiplied as a reference stock.

Extensive phenotyping of plant morphology, seed quality and disease resistance has been carried out on a subset of the faba bean lines, in parallel with the development of markers and genomic tools that are being made publically available. This provides an excellent resource for future genetic analysis, utilising marker assisted selection in breeding programmes. The extent to which genetic knowledge can be transferred among pea, faba bean and *Medicago truncatula* is apparent; for the last, there is an updated

finished genome sequence with associated search and gene expression tools.

Objective 4: management & communication. PCGIN acts as the umbrella organisation for UK legume crop genetics. It has established an excellent network with breeder and end-user communities, and allows constant and useful communication between the UK research base and EU and wider industrial connections. The open structure of PCGIN has facilitated the development of associated projects with a number of industries, including LINK, TSB and EU partnership projects involving both pea and faba bean. The foundation work of PCGIN has provided pulse breeders with the tools to advance breeding programmes in both peas and beans. The enabling technologies which PCGIN has created mean that breeders can improve traits efficiently, creating better uptake and potential for UK pulse crops.

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 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 Exchange).

Scientific (1 - 3) and communication (4) objectives

The scientific objectives of the network are listed below. These objectives relate to the genetic improvement of both pea and faba bean legume (pulse) crops, within a UK context but availing of resources and collaborations within the rest of Europe and world-wide, as and when appropriate.

1. Trait biology, to understand the genetic control of key traits relevant to plant performance and seed quality. This will be based on detailed phenotypic analysis under field conditions of pea populations segregating for economically important breeder priority traits. Linking these traits to genetic maps as quantitative trait loci with adjacent marker and primer information will provide tools directly to breeding programmes. The outputs will have application to related pulse species, where markers may be transferred in consultation with stakeholders.

1.1 To understand the genetic control of yield component traits using PCGIN RIL populations
1.2 To generate RIL populations as a resource for the study of key traits in vining pea
1.3 To understand the genetic control of key seed quality traits using PCGIN RIL populations

2. Genetic mapping, to provide reference maps for marker analysis of the traits under investigation in 1, and additionally support the selection of lines carrying desired traits in the wide crosses established for pea; to build on international links and resources to provide for genetic mapping in faba bean for UK benefit.

2.1 To provide genetic maps for PCGIN RIL populations to underpin QTL analyses
2.2 To provide for genetic marker development in pea
2.3 To provide for genetic marker development in faba bean
2.4 To provide updates on genetic marker development in lupin

3. Genetic resources, to expand the available resources for pea and faba bean, exploiting collaborations with European platforms for pea, and international collections of faba bean.

3.1 To provide novel genetic resources for pea
3.1.1 To obtain novel mutants associated with quality traits
3.1.2 To obtain novel mutants associated with performance traits
3.1.3 To generate NILs for QTL identified by areas 1 & 2
3.1.4 To analyse fast neutron deletions affecting plant architecture

3.2 To expand the genetic resources for faba bean in UK
3.2.1 To establish faba bean inbred lines
3.2.2 To provide phenotypic descriptors of faba bean lines and accessions
3.2.3 To establish mapping populations for faba bean

4. Management & communication, to provide for communication channels, based on established PCGIN communication networks with breeder and end-user communities, and exploiting EU and wider industrial connections.

4.1 To manage PCGIN in a responsive manner
4.1.1 To establish related programmes of work

4.2 To integrate PCGIN with international activities

4.3 To disseminate and publish results

Progress and scientific outputs: results

The sections numbered 1 – 4 below provide a summary and highlights of the research and outputs of PCGIN. The objectives either have been met or superseded and, despite the challenges posed by field trial problems as well as by environmental challenges in 2012, all partners showed a willingness and determination to overcome obstacles and repeat field experiments were carried out in 2013. The primary outputs from every objective are discussed in the relevant sections below. In overall summary, the

network has delivered: a unique set of resources for UK pulse crops, genetic tools and knowledge that are being exploited to underpin genetic improvement of both crops, and fostered a unique and unparalleled set of associated projects with industry that are beginning to capitalise on genetic diversity, and the developed tools and knowledge that facilitate crop improvement. Most importantly, the network has provided for cohesion of the needs of breeding, end user and agricultural industries with the objectives being addressed by the research base, such that the relevance of science to modern breeding objectives has been apparent throughout. The industries involved have stated their approval for both the science and its communication, and all parties have gained benefit from a two-way knowledge exchange. Training was provided to industry staff, through associated projects and through the joint JIC/UEA M.Sc. programme.

In the following sections, the main results and findings are discussed in relation to their relevance and implications. Where methodology is not provided under the results headings, see section 5 for further detail.

1.1 Genetic control of yield component traits

Three sets of recombinant inbred lines (RILs) were developed within PCGIN, using commercial pea cultivars as parents in two-way crosses, giving rise to the following populations that were used for field trials in 2010 – 2013: Brutus x Enigma (BE/EB), Brutus x Kahuna (BK/KB) and Enigma x Kahuna (EK/KE). Seeds were bulked from F_6 RILs, for which genetic maps were developed to enable quantitative traits to be mapped. The three populations were grown in field trials at NIAB and PGRO for agronomic assessment of yield and yield-related traits, including disease and lodging. The 2010 trials were of single plots of every line, whereas those in subsequent years were of triplicate plots of every line. Genetic analysis of the agronomic data from the 2011 trials provided putative quantitative trait locus (QTL) data from interval mapping, which validated some of the data obtained from the preliminary 2010 (single plot) trial data. This included major QTL for mean seed weight (seed size), total yield and components of lodging. A further year's trials (in 2013) were planned to test the validity of these findings.

The 2012 trials and their associated data proved very difficult due to a number of problems at both the PGRO and NIAB sites, exacerbated by the foul weather experienced throughout the UK in that season. The data produced and their limitations were discussed at the PCGIN management meetings in November 2012 and February 2013. The NIAB statistical over-year analysis of 2011 and 2012 data from both PGRO and NIAB sites led to the definitive conclusion that the 2012 field trials should be repeated in their entirety with additional safeguards put in place to ensure that the non-weather components of the 2012 trial problems were prevented. It was still possible to analyse part of the NIAB plot data, adjusted for the negative effect of black-grass on some plots, whereas the more widespread damage of plots at Thorney (PGRO) invalidated much of the analyses. Table 1 (see end of report, p. 23) summarises the significant QTL and associated genetic traits, of which some key data are discussed here.

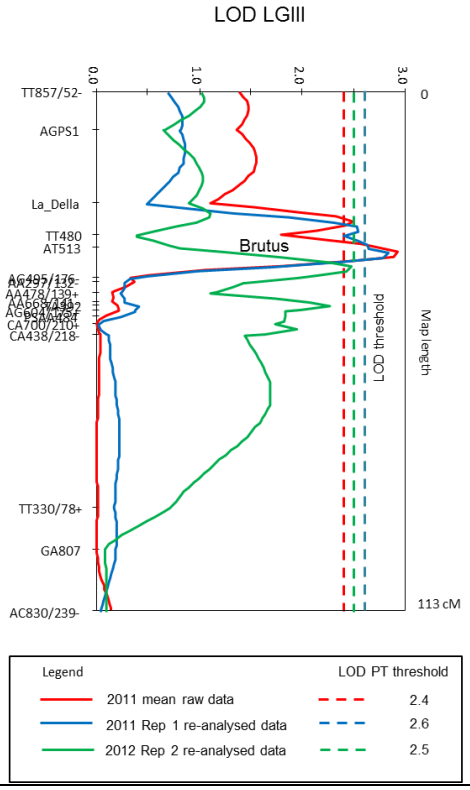

Figure 1: Comparison of significant QTL for yield: KxB RILs, 2011 and 2012 yield data, using interval mapping. Re-analysed data refer to data adjusted by incomplete block analysis. LOD thresholds indicate significant peaks according to permutation tests

Based on adjusted data for BK/KB and BE/EB populations at NIAB, QTL analyses provided confirmation of some earlier data, including significant control of yield associated with markers on linkage groups (LG) I and III. In BK/KB RILs, the LG III QTL for yield shows a major peak of significance in 2011 and 2012 data near the marker AT513. Here the allele from the cultivar Brutus is contributing positively to the genetic effect (Fig. 1). A LG I QTL for yield was apparent in the data acquired for yield over 2011 and 2012 for the BE/EB population (Fig. 2). This QTL was highly significant (mean LOD ~5.0 for mean 2011 and 2012

replicates) across an interval spanning approximately 12 cM on LG I (PSAA474 – AG70/SPDS1 – TA942 marker interval). Here the Enigma parent allele contributes the positive genetic effect. Figure 2 shows QTL data based on individual replicates from 2012, alongside mean data for 2011 and the genetic interval on LG I.

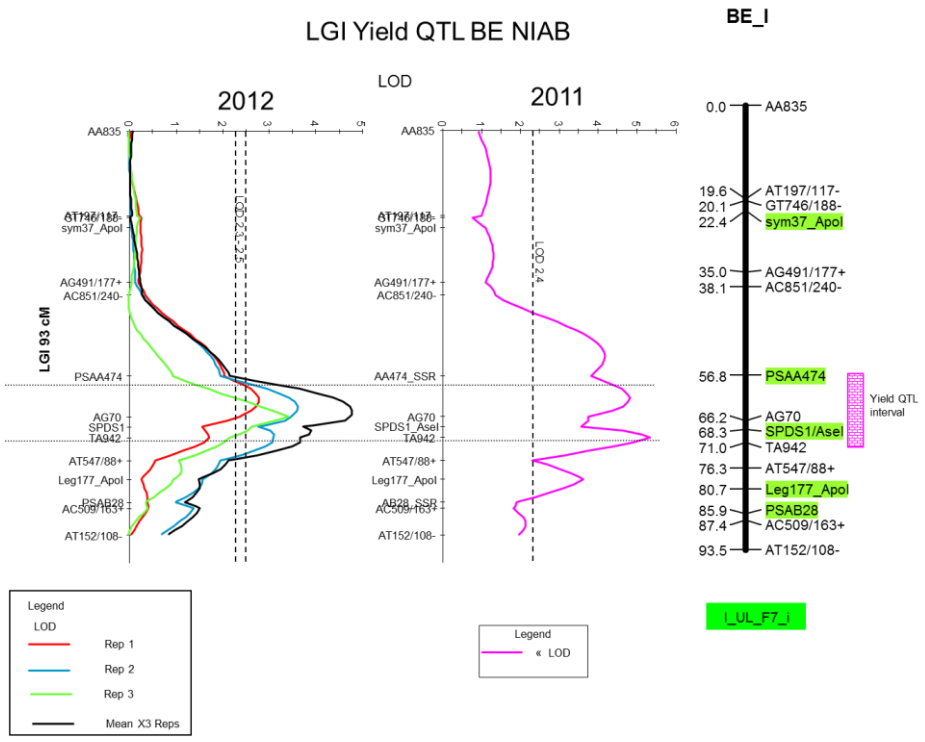

Figure 2: Comparison of significant QTL for yield: BxE RILs, 2011 and 2012 yield data, alongside the genetic marker information relevant to this genetic interval, using interval mapping. Gene-specific markers are highlighted in green. LOD thresholds indicate significant peaks according to permutation tests

The 2013 season provided more stability for the RIL trials at both locations, with the PGRO trials being managed on site at Thornaugh where additional pest control measures were in place. These trials provided reliable data, NIAB statistical analyses were carried out in March 2013, and their genetic analysis is providing confirmation of the major QTL above. Yields at both NIAB and PGRO sites in 2013 ranged from > 3 t/ha to > 5t/ha. A high yielding commercial (control) variety, Prophet, was included at both sites, and gave the highest yields at each, giving a high degree of confidence in the relative ranking order of the RILs from different populations. For each population, several RILs significantly out-yielded both parents and approached the yield of Prophet.

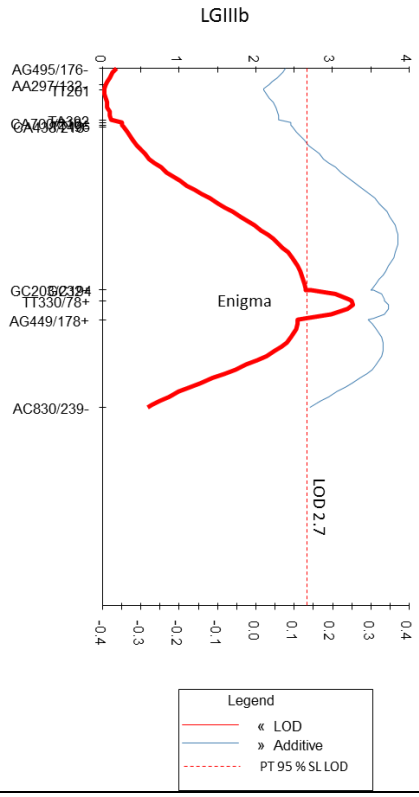

Figure 3: Significant QTL for standing ability in ExK RILs, 2013, using interval mapping. LOD threshold indicates significance according to permutation tests

Our analyses have identified an association between standing ability (lodging) and markers on LG III in the EK/KE RILs (PGRO 2013), in agreement with earlier analyses. Figure 3 indicates that this trait is associated with a similar genetic region as shown for yield determination in the KB/BK RILs in Fig. 1.There are fewer marker data available for the ExK RILs, where Enigma is contributing positively to standing.

Variation in seed size has shown remarkable consistency among

RILs and experiments, including those involving seed colour stability (see later). In all populations involving the cultivars Kahuna or Princess, a major and highly significant QTL for seed size (mean seed weight) is evident on LG I. Additional QTL with lesser effects are evident on LG IV and V (Fig. 4). In all cases, the marrowfat-seeded parent contributes positively to the trait. In populations that do not have such a parent (BE/EB), a QTL for seed size was apparent that co-localised with that for yield on LG I (Fig. 2), with Enigma contributing positively to both traits.

The Brutus x Enigma (BE), Brutus x Kahuna (BK) and Enigma x Kahuna (EK) RILs were taken through to F_{12} seed during 2012 by single seed descent (SSD), with 200, 205 and 178 lines, respectively, and some further catch-up lines in every case. Bulking of F_{12} RIL seeds will be needed for general distribution and future trials. Discussion within the PCGIN management group led to the agreement that the SSD lines be managed in the future by the JIC Germplasm Resources Unit, from which the populations and their genetic map information would be available as a combined resource. In 2013/14, two of the populations are undergoing multiplication as a genetic resource.

Although downy mildew disease did not feature highly during any of the field trials, within PCGIN a source of genetic resistance was identified in a wild *Pisum* accession, JI 15. RILs from the mapping population, JI 15 x JI 1194, were identified which carried the resistance gene along with a short plant stature, wrinkled seed and white flower phenotypes; seeds from these lines have been adopted into a breeding programme to introgress the resistance gene into commercial lines.

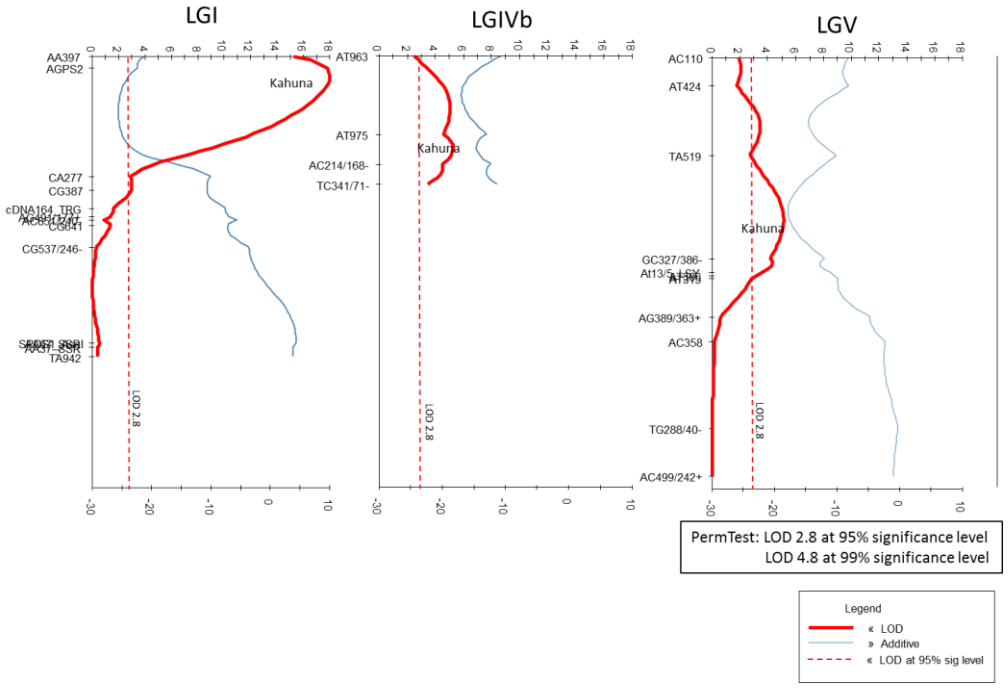

Figure 4: Significant QTL for mean seed weight in ExK RILs, 2013, using interval mapping. LOD thresholds indicate significance according to permutation tests

1.2 RIL populations as a resource for the study of key traits in vining pea

The three-way crosses between Avola, Cabree and Waverex were used to develop RILs for the study of vining seed traits. These populations are available as SSD lines at F7. The parents have been assessed within the LINK project (QDiPS) for sensory traits by the Campden BRI team and a report is available as an output from that project.

1.3 Genetic control of key seed quality traits

1.3.1 Genetic control of seed colour stability

A mapping population was established within PCGIN for the study of seed traits, using a commercial cultivar and an exotic line as parents. At F6, a genetic map was developed based on 178 RILs, and progeny seeds were multiplied for field trials through collaboration with University of Saskatoon (USASK),

Canada, who routinely perform multiplications in Canada followed by a winter nursery in Arizona. This provided sufficient seeds for replicated trials of the population in 2010, 2011 and 2012. Data were collected for plant traits, mean seed weight, seed and cotyledon colour for genetic analyses.

QTL analysis (interval mapping as used for yield above) identified a number of QTL associated with seed traits. Of these, a very significant QTL for cotyledon colour was found on LG I, and a more minor QTL in some but not all experiments on LG III (Fig. 5). Candidate genes were sought within the genetic intervals identified by the QTL analysis.

On linkage group I, a gene encoding a regulator of chlorophyll degradation was chosen for detailed study of its effects on seed colour stability. An in depth analysis of allelic variation in *Sgr* genes was based on sequencing *Sgr* from a selection of lines, including the parents of the RILs, variant germplasm and a selection of other commercial lines. This allowed the classification of genes into a number of groups which all differ by very distinctive features. Most notably, the Sgr regulator, which promotes the loss of green colour from leaves and seeds, is encoded by a gene which is potentially functional in all commercial lines of pea so far examined. In all green cotyledon commercial lines tested, one gene variant is present which is somewhat impaired in its function but it is not disabled. The same is true of two further variants in germplasm accessions where there is potential for transcription of a full-length mRNA and translation to yield a functional protein; sequences that distinguish these variants have been used to develop a suite of diagnostic markers that will be applicable and valuable to breeding programmes. Examples of amplicons generated by primers that distinguish different alleles are shown (Fig. 6).

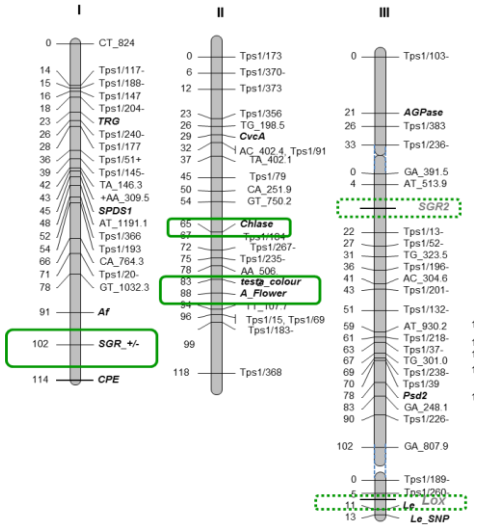

Figure 5: Genetic map for Princess x JI 185 RILs: LG I – III with genes or QTL associated with seed colour indicated in green boxes. Dotted outlines indicate genes which do not function immediately within the chlorophyll degradation pathway

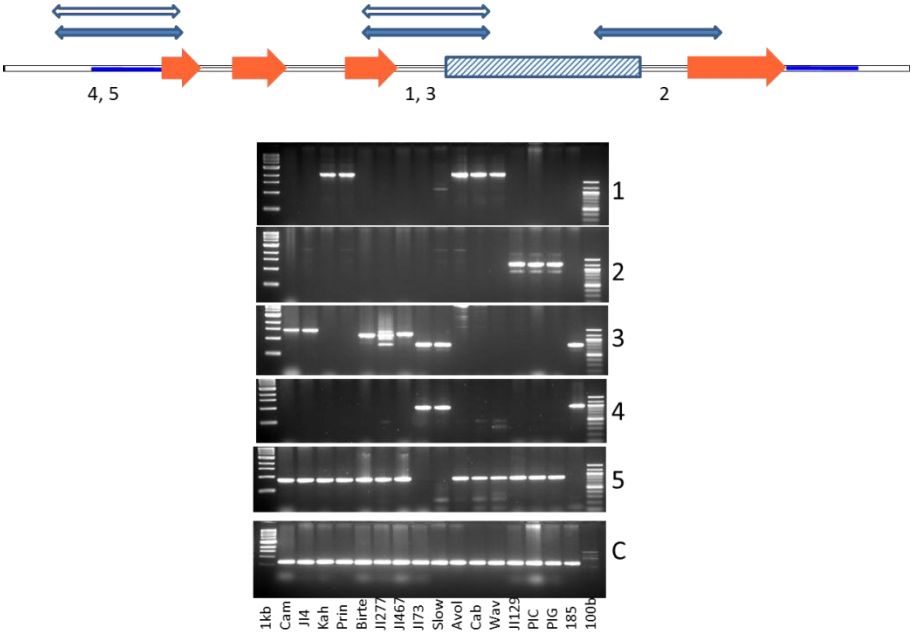

Figure 6. Schematic showing four exons (red arrows) and positions of amplicons 1-5 above, which classify allelic variants of Sgr *in 16 pea genotypes. The hatched box indicates the hyper-variable region within intron 3; the solid blue lines indicate the ends of transcribed regions. Intron 3 variants are shown in panels 1-3 and promoter variants in panels 4 and 5. A control amplicon for all genotypes is shown in panel 6. (Negative control not shown). The markers are 1kb (left) or 100bp (right)*

We have tested a range of *Sgr* and *Sgr2* (a second gene similar to *Sgr*) genes for their effects on loss of chlorophyll, using a transient leaf assay in *Nicotiana benthamiana* (Fig. 7). This system, alongside the identification of other genes involved in loss of pigment, is allowing us to understand the control of the trait and how it may be improved in the crop. The combined knowledge has suggested a strategy for improving colour stability in seeds, namely the generation or identification of a completely non-functional regulator.

Non-functional or null *Sgr* variants have been sought by screening populations of three types: TILLING platform developed from EMS-mutagenised lines (INRA-Evry), fast neutron mutagenised population developed within PCGIN and JIC *Pisum* germplasm. A set of TILLING mutants, including two having an early stop codon are being characterised and crossed into a number of backgrounds to determine their effects on seed and plant traits. Germplasm screens have identified a novel variant, a likely null, where an insertion into exon 4 coding sequence (Fig. 8) is predicted to prevent translation of a functional protein.

These experiments provide new sources of genes that impact directly on a seed trait of economic importance, they illustrate the value of the approaches taken, and are expected to lead to a key publication.

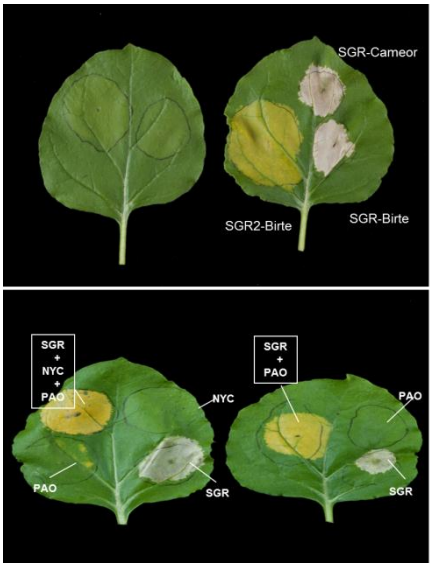

Figure 7: Example of leaves from a whole plant transient leaf assay of genes alone or in combinations in Nicotiana benthamiana. *Top panel: leaf infiltrated with an empty vector (control, left); right: leaf infiltrated with three different* Sgr *genes, showing levels of colour loss. Lower panel: effects of different genes alone or in combination on colour loss; NYC, chlorophyll* b *reductase; PAO, pheophorbide* a *oxygenase*

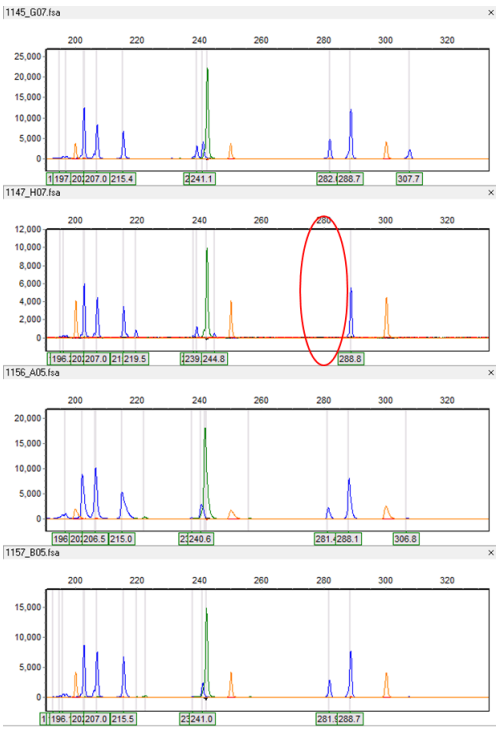

Figure 8: Multiplex fluorescent screen of a set of variant Pisum *germplasm for a number of genes: the predicted amplicon for Sgr exon 4 is present at 279-282 bp in three lines shown but is absent from the line shown with highlighted red circle*

1.3.2 Genetic control of seed nutritional quality

Mutants for the pea seed trypsin/chymotrypsin inhibitor *TI1* gene were obtained from the TILLING platform (INRA Evry) and the effects of the mutations studied. One mutation (C77Y) results in the complete loss of activity of the corresponding protein, leading to a reduction of more than 50% in inhibitor activity compared with wild type (Fig. 9, left pair of bars), where residual activity can be attributed to a second related protein, TI2. These data illustrate the power of TILLING for seed improvement (as above) by delivering a non-GM genotype that differs by a single base pair from its parent; this is especially true in cases where germplasm screening does not yield extreme variants. Analysis of the set of lines shown in Fig. 9 is contributing to a paper in preparation (Clemente et al.), an abstract of which is provided as an annexe to this report.

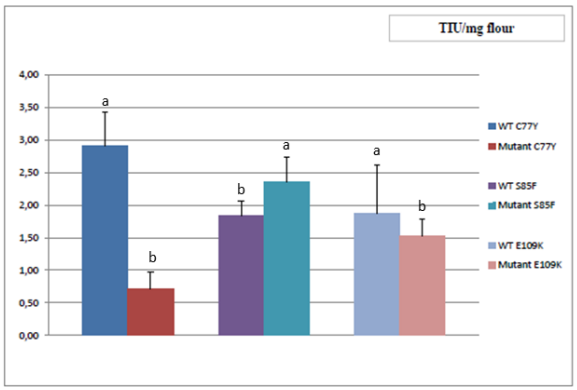

Figure 9. Trypsin inhibitory activities per mg meal of three TILLING mutants (C77Y, S85F, E109K) and their corresponding wild type pea lines. Significant differences ($p < 0.05$) between wild type and mutant lines within each pair are denoted (a, b)

Combining this mutation with other seed protein variants is a powerful strategy for quality improvement. Additional variants include a natural (null) variant for pea albumin 2 (Vigeolas et al., 2008) and several additional nulls that were identified from screening the fast neutron-mutagenised *Pisum* PCGIN resource (Domoney et al., 2013). These mutations in various combinations will allow seed composition to be radically changed, to improve both seed protein quality and quantity. (The variants available through the QDiPS LINK project add a further dimension to these experiments). The impact of such changes on crop yield is being followed up within ProtYield (TSB); here the *Pisum* technologies are being transferred to the faba bean's industry breeding strategies and supporting tool development within that project.

2. Genetic mapping

2.1 To provide genetic maps for PCGIN RIL populations

The development of genetic maps to underpin QTL analyses in the Brutus x Enigma (BE/EB), Brutus x Kahuna (BK/KB), Enigma x Kahuna (EK/KE) and Princess x JI 185 populations is described above. Representative regions of relevant linkage groups are shown in Figs. 2 and 5. It is notable that marker development is much more difficult for the first three populations, where genetic diversity is limited reflecting the cultivated background. For some linkage groups in these crosses, few markers are available. This has the advantage of limiting the regions of linkage groups that can be used to explain genetic variation of traits. For the wider cross, more markers are available (Fig. 5), where the wild background of JI 185 contributes much variation for both molecular markers and traits.

We expect that the PCGIN RIL mapping information (and see below) will be made available, along with the genetic stocks, as a resource for the wider community.

2.2 To provide for genetic marker development in <u>pea</u>

A selection of PCGIN RILs was made so that PCGIN would participate in a large genome wide association study (GWAS) being coordinated by USASK (University of Saskatoon). A set of lines having contrasting genotypes was selected from the three PCGIN populations and used for field trial multiplication at USASK in 2012 with some preliminary trait analysis carried out. Fig. 10 shows an analysis of the BK/KB population, based on the numbers of shared markers, where lines representing the diversity of marker groups were selected (highlighted); this analysis was repeated for the other two populations (not shown) and 39 RILs plus three parent lines were included in the GWAS. The multiplied lines are scheduled to be distributed among a large number of field sites for trait analysis and the trait analysis used linked to genome wide marker data. The GWAS will be based on genotyping approximately 960 lines in total, using a 1536 SNP panel of single nucleotide polymorphisms (SNPs) in an Illumina GoldenGate assay. This will

indicate the potential for enriching the map information that supports the QTL analysis of three whole populations.

Initial data from USASK indicate that some PCGIN RILs show outstanding performance in the field; the wider use of some of these RILs in breeding programmes is being debated within PCGIN.

2.3 To provide for genetic marker development in <u>faba bean</u>

We have applied next-generation sequencing technology to screen for polymorphisms to develop molecular markers for mapping traits of interest in *V. faba*. Using 454 sequencing (Solexa) in two diverse *V. faba* lines, Albus and BPL10, we have obtained transcriptome sequence suitable for performing *de novo* assembly and SNP discovery. 14K gene transcripts common to both BPL10/Albus were aligned against *Medicago truncatula* to identify 40,000 putative SNPs. 888 of these were subsequently developed into KASPar assays and validated across 33 diverse, in-bred *V. faba* reference lines. 726 assays were observed to be highly reproducible across the diverse faba germplasm and have been used to produce a series of new genetic linkage maps.

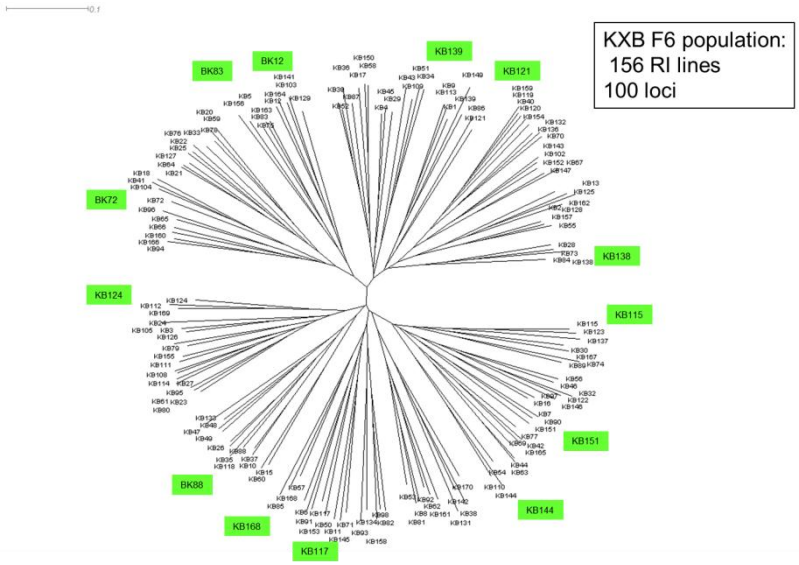

Figure 10. Selection of RILs from the BK/KB populations for GWAS, based on marker diversity. Groups of lines having shared markers are indicated and RILs chosen to represent the marker groups highlighted in green

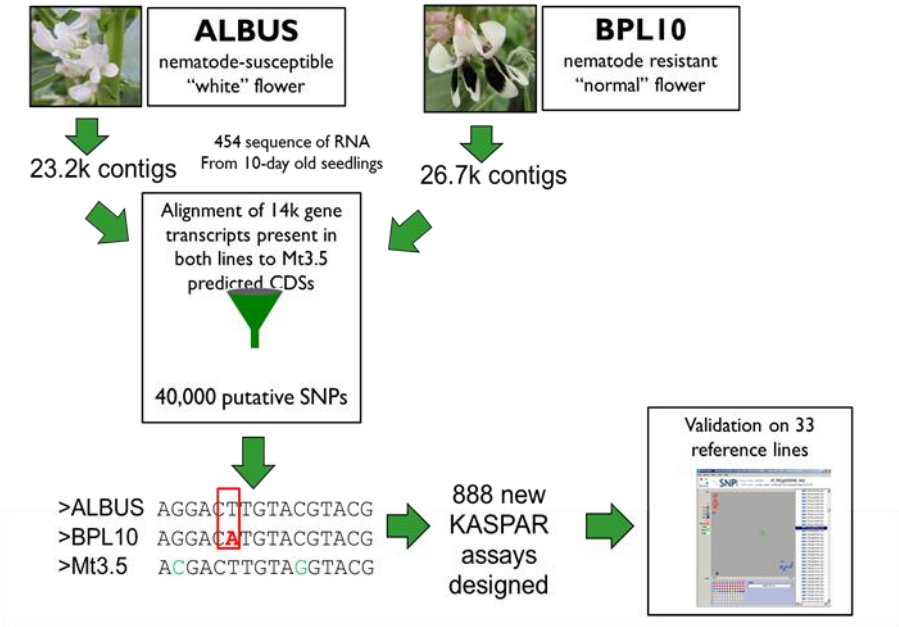

Figure 11. Strategy for genetic marker development in faba bean

3. Genetic resources, to expand the available resources for pea and faba bean

3.1 To provide novel genetic resources for <u>pea</u>

Table 2 (see end of report, p. 23/24) summarises the PCGIN expanded genetic resources for pea and faba bean, some of which are discussed here. We have exploited variation within natural germplasm, as well as that identified within mutagenised populations to identify, isolate and characterise novel alleles of genes which, when introduced into appropriate backgrounds, facilitate meaningful comparisons with wild types. Examples include the pea albumin 2 seed protein variant, identified in an exotic background (Vigeolas et al., 2008), which was backcrossed three times to a cultivar. The parent and mutant lines are currently being multiplied in UK and New Zealand to support feeding trials as part of the ProtYield project. Detailed analysis of the consequences of such a mutation and its consequences for seed quality has proceeded within PCGIN. The PCGIN 'high-yielding' RILs, identified on the basis of 2011/12 field trial comparisons, are being used for the introgression of the seed protein variants to provide ultimately a set of near-isolines for field trials.

Further examples of novelty include the fast neutron-derived mutations (Domoney et al., 2013) and the TILLING alleles, as described above for genes encoding seed inhibitors and seed colour stability. These mutants provide non-GM solutions to plant and seed trait improvement, which should have wide acceptability. The marker information developed for the variant genes in every case will enable marker-assisted breeding following introgression of novel alleles. In some cases, these markers will be 'perfect' but in others (for example, the downy mildew resistance gene(s) in JI 15), further work is needed to develop markers based on candidate or closely linked gene. The availability of a number of genome sequences for legume model and crop species is greatly facilitating the identification of candidate genes, based on overall synteny of genetic maps among species.

3.2 To expand the genetic resources for faba bean in UK

NIAB has acquired a collection in excess of 900 *Vicia faba* accessions comprising a broad range of UK, EU and Middle Eastern cultivars and landraces. Seed numbers per accession vary considerably, but provide enough starting material for further development. Seeds are currently held in the Long Term Seed Store at NIAB.

3.2.1 To establish faba bean inbred lines

More than 130 accessions have been entered into a single seed descent (SSD) programme to obtain inbred germplasm. Currently 89 accessions have been taken to ≥SSD6 and many more lines continue to be in-bred to SSD6 and beyond. Seventy-five genome-wide KASPar markers have been developed from CAPS/SNAPshot markers derived from *V. faba* lines Vf6 and Vf27 for screening genetic diversity (Cottage *et al.,* 2012). A panel of 39 robust assays have been applied to subsequent generations of SSD lines to track the progressive reduction in heterogeneity within the programme. The toolkit presents an excellent resource for determining levels of hybridity/genetic diversity within faba accessions. To date 89 faba accessions currently at ≥SSD6 have been screened for heterogeneity. The majority are now observed to be homozygous for all 39 markers tested. Homozygous markers from the panel were included in each screen as a quality control to ensure no cross-pollination had occurred. The remaining lines will be harvested at SSD6 and genotyped imminently.

A highly inbred line of Hedin (over 30 generations of selfing) was obtained from Wolfgang Link (University of Gottingen) and has been further multiplied. This is proposed as a reference faba bean genotype and currently 3600 seeds are available.

3.2.2 To provide phenotypic descriptors of faba bean lines and accessions

Phenotyping/morphological data for the traits listed below have been compiled in spreadsheet format. A total of 70 lines at SSD6 have been fully phenotyped, and a further 81 have been partially genotyped to date.

Assessments at Flowering time:

Number of branches, stem colour, leaf size (taken as two separate measurements, length and width, from a random leaf at 2nd flowering node), stem width cm at the base of the stem, height (taken from soil level to the growth tip of the plant), height to first flower, flower colour, number of nodes to first flower, number of flowers at 1st, 2nd and 3rd flowering node, date of first flower

Assessment at maturity (senescence):

Height, number of branches, stem width, height to first pod, number of nodes to first pod, number of pods and seeds per pod at 1st, 2nd and 3rd flowering node, number of seed harvested from a main stem and branches, total number of seeds, date of maturity

Seed assessments:

Colour, size, weight, shape, type and hilum colour

Disease:

Methods have been established for phenotyping for downy mildew (*Peronospora viciae)* and rust (*Uromyces fabae)* using a percentage coverage scale (%). A range of the accessions from the NIAB collection were screened for resistance to the two pathogens (70 accessions for downy mildew, 86 for rust) in two week old plants in inoculated growth room trials. A number of lines tested demonstrated a relatively high level of resistance to DM (NV13 and NV 38) and will be used in future studies to identify loci contributing to improved disease resistance to the pathogen. For rust none of the varieties screened exhibited greater resistance than the 'Icarus' control. A chocolate spot screen was carried out in 2013 using internationally defined methods. However, none of the entries proved to have a high level of resistance, including lines which had exhibited some degree of partial resistance in other countries.

3.2.3 To establish mapping populations for faba bean

Two new faba bean mapping populations (F2:F3) have been developed from crosses between BPL10 x Albus and 29H x Albus. The two populations have since acted as a foundation for the TSB funded StemNem project that aims to identify sources of genetic resistance to *D. gigas.* RNA sequencing data derived from BPL10 and Albus have now been used to develop ~800 new KASPar assays. The KASPar assays developed under PCGIN were deployed to screen and generate individual genetic maps for four, bi-parental F2:F3 faba populations (NV643 x NV648 (segregating for white flower/Stem Nematode resistance), NV657 x NV643 (white flower/Stem Nematode resistance), NV644 x NV153 (Dwarf/white flower) and NV639 x NV658 (closed flower/open flower)) A consensus map, comprised of 643 common markers, was constructed (Fig. 12) and consisted of six linkage groups with a total length of 1,596 cM (O'Sullivan *et al.*, in preparation). A further population from a cross between downy mildew susceptible and resistant lines has been initiated in the early part of 2014.

The degree of co-linearity between the sequenced regions of the *V. faba* transcriptome and corresponding orthologous regions in *M. truncatula* was assessed using BLAST and comparative mapping (Strudel, http://ics.hutton.ac.uk/strudel/); Large regions of the *V. faba* genome were observed to demonstrate macro-synteny with *M. truncatula* (Fig. 13). In general there was a high degree of conservation in gene order over few, extensive blocks of co-linearity. For example linkage groups Vf2 and Vf3 demonstrated the greatest levels of synteny with Mt3 and Mt1, respectively (Fig. 14).

To demonstrate the power of synteny-based mapping we utilised the new resources to map one of the loci responsible for white flower colour and zero tannin phenotype (Zt-1) using the NV643 x NV648 population (Webb *et al.*, in preparation). Orthologous loci have already been identified in *M. truncatula* and *P. sativum* (Transparent Testa Glabra 1 (TTG1)/A2, respectively) (Pang *et al*., 2009, Hellens *et al.,* 2010). The F2 generation of the NV643 x NV648 population was genotyped and the F3 generation scored for flower colour (and StemNem resistance). A QTL analysis identified a peak between markers Vf_Mt3g92810_001 and Vf_Mt3g094760_001 and gene orthologs of the Medicago TTG1 (Mt3g092830, Mt3g092840), were identified within this interval. The coding region of *Vf*TTG1 was cloned and sequenced from NV648 (coloured flower); however a complete *Vf*TTG1 sequence could not be obtained in the white-flowered NV643 line. This indicated a difference in the coding sequence or a deletion, compared to NV648. Genome walking was utilised to obtain the sequence of the 5' region of the *Vf*TTG1 locus in NV643. This identified a deletion spanning from the middle of *Vf*TTG1 upstream of promoter elements. This strongly indicated that a deletion in the coding region of *Vf*TTG1 in *V. faba* was responsible for the loss of floral pigmentation in NV643.

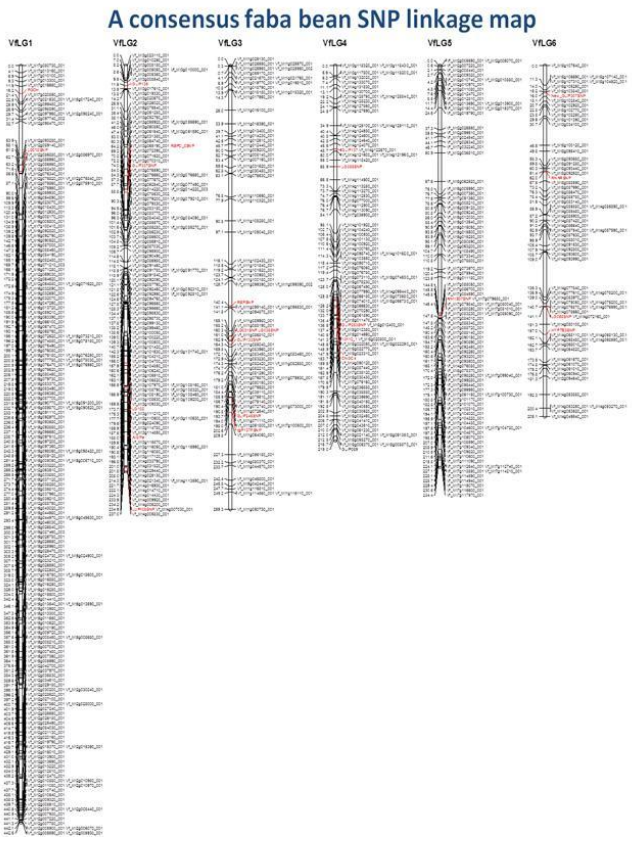

Figure 12: A consensus linkage map for faba bean. (Information on markers is available on request)

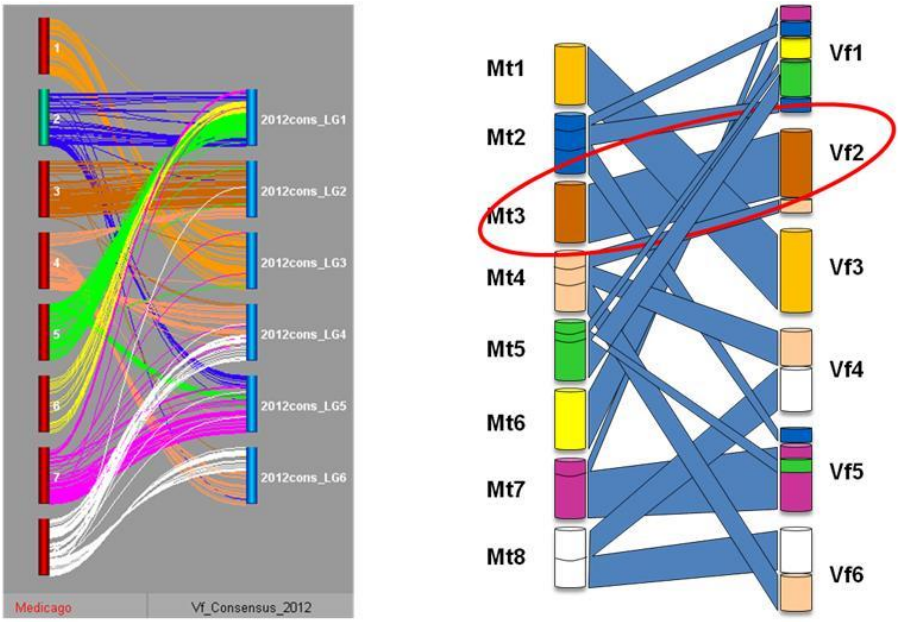

Figure 13: Synteny mapping using Strudel software

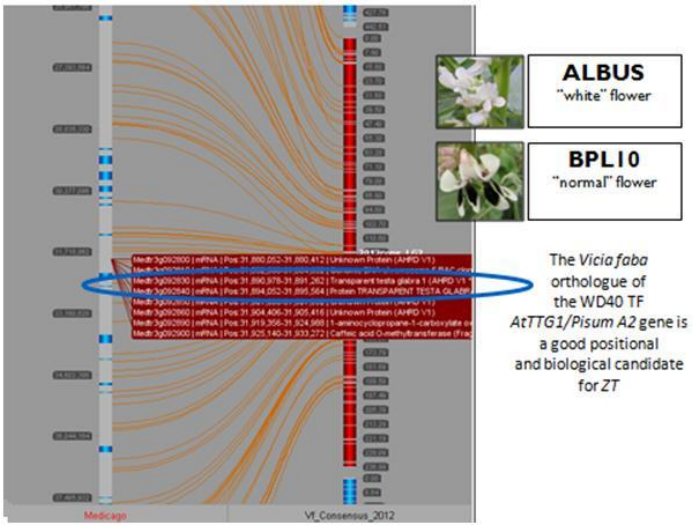

Figure 14: Identification of putative gene for white flower/zero tannin in faba bean

4. Management & communication

The network provides a vital link between academic science at JIC and NIAB, applied science at PGRO and the wider industry, both breeders and end users. All partners will testify to very open communication channels, the 'open door' policy at JIC and NIAB for partners, stakeholders and associated visitors from within UK and abroad. PCGIN has extended its reach into industry since its inception and has hosted several meetings involving many disparate groups. (Examples include meetings involving participants from BEPA to a Chinese industry with an interest in end-use traits and UK pulse resources, farmer groups and supermarkets).

4.1 To manage PCGIN in a responsive manner

4.1.1 To establish related programmes of work

We have established communication networks within PCGIN which include breeder and end-user communities. These have been developed to initiate new projects, including those under BBSRC/Defra LINK, TSB and more recently HDC. The related programmes have been built upon PCGIN as a platform and exploit resources developed within PCGIN. The related projects include as partners several industries that are relevant to the broad uses of both pea and faba bean as food and feed, ranging from canning, vining and picking vegetable industries, fish and animal feed industries, together with supermarkets having an interest in promotion of both 'home-grown' and novelty.

PCGIN-fostered projects include:

QDiPS (BBSRC/Defra LINK);
Bruchid LINK (Defra);
ProtYield (TSB);
OptiBean (TSB);
Stem Nem resist (TSB);
GWAS consortium (USASK);
Downy Mildew resistance (HDC);
Cambridge University PhD project 'Improving the pollination of field bean (*Vicia faba*) - how do floral characteristics affect pollinator behaviour?' utilises inbred lines from NIAB *Vicia* collection

A recent success is a DRINC2 (BBSRC) project, which will utilise PCGIN/QDiPS resources in a (JIC/IFR/ICL) collaboration; PCGIN industry contacts and materials will be used in a second DRINC2 project (UEA, JIC)

In the pipeline are:

BBSRC/India initiative (two legume projects, one on *Pisum*, and one on *Lathyrus*);
ERA-CAPS - Faba bean Adaptation, Biodiversity in Landraces, Evolution and Domestication: consilience between genomics and history to optimise faba bean's role in agro-ecosystems

4.2 To integrate PCGIN with international activities

PCGIN has been engaged with the wider EU activities since its inception. This engagement has facilitated the exploitation of, for example, the TILLING platform for pea genomics and the international faba bean work in Spain. The FP7 LEGATO (LEGumes for the Agriculture of TOmorrow) project which has recently begun (January 2014) will use some resources and tools developed within the QDiPS LINK project. The wider PCGIN will benefit from engagement with LEGATO, ensured by participation of JIC and PGRO. BBSRC funding supported an international workshop on Orobanche control "Building a new research alliance to reclaim faba bean production area abandoned to *Orobanche*", Morocco, October 2013. Follow-on work has been identified and further funding will be sought.

PCGIN research activities have been supported so far by engagement internationally with ICARDA scientists and representatives from North Africa, USASK (Canada), INRA-Evry and INRA-Dijon (France), CSIC, Spain and, more recently ICRISAT, India. Engagement with ICARDA scientists has resulted in collaboration on the ICARDA chocolate spot nursery which was run in the UK in 2013, exploring reaction of selected accessions and PCGIN lines to UK national *Botrytis fabae* populations (97 isolates collected in 2012). A visiting researcher from INAT, Tunisia, has worked on a baseline evaluation of vicine and convicine levels in sets of faba bean lines from PCGIN and elsewhere, and ongoing collaboration with ICARDA is developing a "MAGIC" population for faba bean which will be available as a potential resource.

4.3 To disseminate and publish results

Stakeholder interactions have occurred formally through involvement with PGRO open days, at which PCGIN objectives, research and outputs have been presented and publicised. Various press articles and coverage document this. Beyond this, a wide variety of talks have been presented to very diverse audiences, ranging from industries with little understanding of genetics through to international conferences devoted to legume genomics and genetics. Demonstrations have also been presented at the Cereals Events, NIAB Open Days and at NIAB Innovation Farm events.

Included here are presentations to Wherry, Limagrain, Germains, Agro-Seed Services (Belgium), Premier/Princes, BEPA, Campden BRI, NFU, Alpha Farmer Group, Mardlers farming group, Defra visiting groups, the JIC/SL annual science meeting, and the ICLGG (International Conference on Legume Genetics and Genomics), IFLRC (International Food Legumes Research Conference) and NAPIA (North American Pulse Improvement Association) as examples of international conferences.

5. Materials and methods

5.1 Plant materials Recombinant inbred lines of pea were developed by single seed descent (SSD) from the cultivars Brutus (B), Kahuna (K) and Enigma (E) in three-way crosses, and from Princess and JI 185. In all cases, seeds were bulked from F6 lines for which DNA was extracted and genetic maps were established. In the case of Princess x JI 185, F7 seeds were bulked to F8 in Canada, to F9 in an Arizona winter nursery, to F10 in two Canadian locations (SPG/Ros) and to F11 (Canada SPG x 2 sites). Seed samples were returned to JIC for phenotyping. The BKE crosses (>200 RILs per population) were bulked to F8 at PGRO (polytunnel), F9 in single plots at NIAB and PGRO, to F10 in replicated trials at NIAB (two populations) and PGRO (one population). These replicated trials were repeated over three consecutive seasons. For the faba bean SSD breeding program and genotyping, over 130 lines were selected for SSD. Plants were selfed and a single seed taken forward to the next generation. DNA was extracted at each subsequent generation and genotyped using 39 KASPAR assays (Cottage *et al.,* 2012). Faba bean mapping populations (F2:F3) were developed from crosses between BPL10 x Albus and 29H x Albus.
Additional plant materials were from the JIC and USDA pea germplasm collections, the TILLING platform (INRA-Evry, France) and a set of *Vicia faba* accessions comprising a broad range of UK, EU and Middle Eastern cultivars and landraces acquired for storage at NIAB.

5.2 Genetic mapping and resources Primer information is available as supplementary material. For pea populations, mapping was based primarily on retroelement markers (Knox et al., 2009) and all gene-specific sequence information available (Aubert et al., 2006; Embl/GenBank or not published). Quantitative trait measurements based on field trials were subjected to incomplete block analysis and adjusted according to standard methods used at NIAB for recommended list trial evaluation. Quantitative trait locus evaluation was performed using MapQTL programmes; data were obtained using Interval Mapping and Kruskal-Wallis analyses, where the latter may handle data that do not show a normal distribution. Statistical analyses were based on permutation tests, LOD scores and significance calculated within the packages. For faba bean, **l**inkage mapping and QTL analysis were conducted using MapDisto 1.7.5.1 for Windows (Excel 2007 version) groups (Lorieux, 2012). This utilised the automap function with the threshold set at a LOD of 5. Synteny mapping between *V. faba* and *M. truncatula* was performed using BLAST and comparative mapping software strudel (http://ics.hutton.ac.uk/strudel)

5.3 Gene sequencing and screening Pea cDNA and gene sequences were amplified according to standard reverse transcriptase (RT)PCR and PCR conditions, with the exception of long PCR amplicons, which were prepared using GoTaq Long polymerase with 10-12 minute extension times. Sequencing was performed by TGAC (Norwich) or by Eurofins (MWG); next generation sequencing of long repetitive DNA and assembly was carried out by MWG. Faba bean RNA was extracted from seven day old seedlings of NV643-4 (Albus) and NV648-1 (BPL10) using a Qiagen RNeasy Plant Mini Kit. Sequencing was performed using Roche 454 GS-FLX Titanium chemistry. Raw reads were assembled into contigs using Newbler v2.3. Multiplex fluorescent screening of a set of variant *Pisum* germplasm for variants was carried out by IDnaGenetics (Norwich).

5.4 Faba bean tool development Alignment and assembly of sequence data contigs were based on coding, whole transcript and protein sequences obtained from the *Medicago truncatula* genome sequencing project (International Medicago Genome Annotation Group). Version 3.5 of the annotated genome, consisting of 8 chromosomes and a number of unanchored BACs, was downloaded from http://www.medicagohapmap.org/downloads.php (01/12/10). Contigs from NV643-4/NV648-1 were aligned to each other and to transcripts of *Medicago truncatula* using GS Reference Mapper 2.3 (Roche 454 Life Sciences). Contigs which gave unique reciprocal best hits over at least 100 bases, at a BLASTn cut-off of E^{-30} in the respective faba bean transcriptome assemblies, and to a single predicted *Medicago truncatula* gene over more than 100 bases, were selected for further analysis.

SNP identification and KASPAR marker development utilised complete contigs of each NV643-4:NV648-1 pair aligned using ClustalW. Alignments containing SNPs more than 50bp from either end of a contig were selected for marker development. For each SNP, a consensus sequence with the target SNP and a minimum of 50bp either side was then screened against the database of *M. truncatula* v3.5 (http://www.medicagohapmap.org/advanced_search_page.php?seq), using the BLASTn function.

Consensus sequences matching multiple chromosomes of *M. truncatula*, those with multiple copies on the same chromosome, or *M. truncatula* alignments which contained introns located closer than 50 bp either side of the SNP were all excluded. Sequences with more than 50 bp between the target SNP and the nearest intron site on the corresponding *M. truncatula* contig were truncated at the nearest intron-exon junctions to avoid primers being designed across intron-exon splice sites which would not anneal to intron-containing sequence. The remaining SNP-containing sequences were then searched in BLASTn against all contigs of NV643-4 and NV648-1. Any SNP-containing alignments with hits to more than one contig

from each line were excluded. A total of 887 KASPAR assays were designed by KBioscience (LGC Genomics, U.K.). These were coded as Vf_MtXgYYYYYY_ZZZ where X is indicates the Medicago chromosome number, YYYYYYY the predicted gene identifier of the reciprocal best BLAST hit *Medicago truncatula* orthologue as given in version Mt3.5 of the Medicago genome assembly (Young et al., 2011), and ZZZ the number of the assay. KASPAR marker assays were validated, using DNA extracted from the leaf of a single seedling from a panel of 33 diverse *V. faba* lines (Fulton et al., 1995).

5.5 Transient and other expression assays In order to predict the effects of individual genes, their mutant alleles and combinations thereof on colour stability, assays were established in *Nicotiana benthamiana* leaves, using methodologies described (Sainsbury et al., 2009). Leaves were harvested between two and seven days after infiltration; infiltrated areas were sampled for chlorophyll *a* and *b* and photosynthetic efficiency measurements. These values were used to determine loss-of-function for enzymes and regulators of the chlorophyll degradation pathway. Assays of trypsin and chymotrypsin inhibitors and their mutants were performed according to published methods (Clemente et al., in preparation).

6. Context and summary PCGIN outputs in relation to forward plan (FP) and FP objectives

6.1 Context for PCGIN

The promise for UK grown pulses continues to be strong but requires promotion on several fronts. The three-crop rule, problems with the control of black grass in cereal crops, poor grower rotational choices in recent years and strong worldwide demand provide strong justification for PCGIN and drivers for its associated projects. Economic analysis by ProCam provides evidence that spring peas, grown for the human consumption market, have given the most consistent performance over the last 10 years, with an average gross margin of £643/ha (Farmers Weekly, August 2013). Within PCGIN, we are contributing to promoting the true economic value of pulse/legume crops, whereby the value (yield benefit) of the nitrogen fixed by the legume crop to the following crop in the rotation and saving in fertiliser cost is rightly attributed to the legume.

Equally, the environmental benefit of legume crops has not been measured correctly in the past, and evidence that biological nitrogen fixation does not produce N_2O (Canfield et al 2010) plus the fact that a legume crop replaces one that would use nitrogen fertilizer, and contribute N_2O emission, must be included within calculated benefits. Nitrogen fertiliser has a high energy demand and its production, distribution and use is the main source of agricultural N_2O, a destructive greenhouse gas. Agriculture contributes 75% of European N_2O emissions, corresponding to between 12 and 25% of total European greenhouse gas production (Crutzen et al., 2008).

The EU continues to be approximately 30% self-sufficient in home-grown protein, largely used for animal feed markets. Concerns that import sources may not always be available in the future have led to heightened interest among feed industries in promoting home-grown sources. Additional pressures on land use and environmental concerns are leading to new models of a future UK agriculture, where legume crop areas are substantially increased (see for example, Zero Carbon Britain 2030: A New Energy Strategy, Centre for Alternative Technology, 2010, eds Kemp, M., Wexler, J., ISBN 978-1-902175-61-4).

There is much growth in the global gluten-free market, where opportunities and technical solutions with pulse (pea) protein are being developed for increasing numbers of consumers, who are exploring gluten-free diets following general health concerns (for example, http://www.cvent.com/events/bridge2food-webinar-cosucra-2014/event-summary-11152cd7bd984d1cb1df44b86fcba684.aspx). Markets for gluten-free protein and novel foods are expanding rapidly (see http://www.bloomberg.com/news/2014-04-23/you-will-eat-your-peas-now-as-big-food-binges-on-protein.html) and opportunities are being exploited by Canadian growers. The United Nations has proclaimed 2016 as the 'International Year of Pulses', which may promote marketing opportunities, similar to those enjoyed by others on receipt of this designation.

For the economic, social and environmental benefits of pulse crops to be realised, the discipline of genetics and development of associated resources based on fundamental knowledge are key. Targets are changing for all crops globally and research needs to respond to targets in a timely manner. The development of genetic resources and associated tools within PCGIN is planned, so that high-throughput screening for new targets will build on established protocols optimised for broad-ranging application. Examples include the printing of DNA plates from the wider pea germplasm, representing exotic, landrace and cultivated genotypes now amenable to rapid genotyping and screening for novel alleles. This and related research on faba bean is providing a wealth of variation for economically important traits, whereby

it is envisaged that a period of experimentation, crossing and validation leads to more applied projects, involving industry and large-scale field trials.

The context and longer term implications of the ground-work delivered by PCGIN continue to be established through extensive discussions with the relevant industries. PCGIN has engaged effectively directly and indirectly (through associated projects) with most of the major stakeholders in UK and others within Europe and world-wide. The network can now extend to others, whose input to UK home and export markets may help to clarify the true value of improved traits to feed and food markets. Attempts to gather reliable market information have proved to be notoriously difficult with estimates being at variance with the experience of industries, and data provided previously by BEPA are no longer available. Discussions with PGRO, British Growers, Frontier, Wherry & Sons, Hutchinsons and Andersons are providing new information, which requires coordination and communication. Interactions with Mack Multiples, HDC and the Legume Industry Panel have and are being exploited in new opportunities for the uptake of resources that are direct or indirect outputs of PCGIN, besides the ongoing projects and interactions listed above (section 4.1.1), where resources have been taken into new projects and consortia.

6.2 Outputs and forward plan

Output 1. Platform that supports UK pulse crop improvement

FP1.1: To maintain a platform that supports wide-ranging projects in UK, EU and world-wide via further development of the websites (www.viciatoolbox.org; www.pcgin.org)

FP1.2: Utilise the technologies available through collaborations (USASK, INRA, ICRISAT, ICARDA) to maximise opportunities for UK legume crop improvement. PCGIN has generated many resources that require further exploitation in order to continue to identify novel sources of genetic variation and for the UK to benefit fully from them.

FP1.3: Disseminate results of scientific research and the means by which they may be exploited by industry

FP1.4: Establish industry-led forum to promote the uptake of newly characterised resources into associated projects (TSB, Agri-Tech; BBSRC LINK where prior support established) (*A/T/I)**

Output 2. Major QTL identified for a number of plant and seed traits in pea, including components of yield, major disease and seed quality traits

FP2.1: Field trials of selected RILs, to represent the BK, BE and KE genetic variation, crosses involving these lines and the novel mutants characterised within QDiPS

FP2.2: Develop enhanced and relevant marker data to support crossing of high-yielding lines and introgression into breeding programmes (*A/T/I*)

FP2.3: Develop gene-specific marker data relevant to the LG I downy mildew resistance locus in the JI 15 derived pea RILs; determine the relationship with other germplasm resistant stocks and UK problem (*A/T/I*)

FP2.4: Develop mapping population from initial faba bean cross between downy mildew resistant and susceptible lines for production of a genetic map and subsequent QTL analysis

FP2.5: Undertake QTL mapping to identify resistance to *Ascochyta fabae* in the faba bean (Inra29H x Albus) population

FP2.6: Access further material claimed to have chocolate spot resistance for evaluation in controlled environment and field screens, create mapping populations. Develop understanding of variation in *Botrytis fabae* to aid interpretation of variable host responses.

Output 3. Novel variants identified that have the potential to change pea seed composition dramatically (reduced or no antinutritional proteins, altered carbohydrate and protein profiles) and prevent loss of colour from seeds for food uses

FP 3.1: Assess the effects of variant genes on 'bleaching' (seed colour stability), introgressing TILLING and variant germplasm alleles, including novel 'green cotyledon' gene (*A/T/I*)

FP 3.2: Classify germplasm variants for a catalogue of crop-relevant genes, using high-throughput screens

FP 3.3: Evaluate the consequences for combinations of mutations on seed composition and plant yield (in parallel with ProtYield, and beyond its lifetime) (*A/T/I*)

Output 4. Toolkits for breeding programmes, including markers and genetic resources

FP4.1: Enable assessment and utilisation of novel genetic resources, including those outputs from QDiPS via SNP platforms, transcriptome resources, synteny with sequenced genomes and candidate gene identification

FP4.2: High-throughput screening of germplasm resources as an industry route to novel gene alleles

FP4.3: Develop an EMS/ENU mutant population in Hedin (reference faba bean) background to generate novel variation for genetic improvement, and establish an efficient transformation system to aid functional analysis of important traits identified in mutant populations

FP4.4: Transcriptome sequencing of faba bean (Hedin) to develop markers linked to traits of interest

FP4.5: Improvement of faba bean genetic maps through addition of SSR markers identified from sequenced faba bean transcriptomes

(*A/T/I)* indicates where opportunity with Agri-tech, TSB or other industry interaction has been discussed. For the last, this may represent, for example, a training opportunity for an industry worker at a research centre. Not listed in outputs is the training of industry that has already taken place, which includes that of Limagrain and Wherry & Sons staff at JIC).*

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.

AUBERT, G., MORIN, J., JACQUIN, F., LORIDON, K., QUILLET, M.C., PETIT, A., RAMEAU, C., LEJEUNE-HÉNAUT, I., HUGUET, T. and BURSTIN, J. (2006) Functional mapping in pea, as an aid to the candidate gene selection and for investigating synteny with the model legume *Medicago truncatula*. Theor Appl Genet **112**, 1024-1041

CANFIELD, D.E., GLAZER, A.N. and FALKOWSKI, P.G. (2010) The evolution and Future of Earth's nitrogen cycle. Science **330**, 192-196. DOI: 10.1126/science.1186120

CHINOY, C., WELHAM, T., TURNER, L. MOREAU, C. and DOMONEY, C. (2011) The genetic control of seed quality traits: effects of allelic variation at the *Tri* and *Vc-2* genetic loci in *Pisum sativum* L. Euphytica **180**, 107-122

CLEMENTE, A., MARÍN-MANZANO M.C., ARQUÉS M.C. and DOMONEY C. (2013) Bowman-Birk inhibitors from legumes: utilisation in disease prevention and therapy. In: Bioactive Food Peptides in Health and Disease. Eds B. Hernández-Ledesma and C.-C. Hsieh. InTech Pub ISBN 978-953-51-0964-8, pp 23-44

CLEMENTE, A., MARÍN-MANZANO, M. C. and DOMONEY, C. (2011) Exploiting natural variation in legume Bowman-Birk inhibitors to dissect their potential role in human health-promoting programmes. Can. J. Plant Sci. **91**, 377-378

CLEMENTE, A., MARÍN-MANZANO, M.C., JIMÉNEZ, E., ARQUÉS, M.C. and DOMONEY, C. (2012) The anti-proliferative effect of TI1B, a major Bowman-Birk isoinhibitor from pea (*Pisum sativum* L.), on HT29 colon cancer cells is mediated through protease inhibition. Brit J Nutr 108, S135-S144

CLEMENTE, A., MARÍN-MANZANO, M. C., JIMÉNEZ, E., MORENO, F. J., RUBIO, L. A. and DOMONEY, C. (2010) The anti-proliferative properties of Bowman-Birk proteins from legumes on colon cancer cells are related to their ability to inhibit trypsin- and chymotrypsin-like proteases. In: Abstracts of the BIT Inaugural Symposium on Enzymes & Biocatalysts, Shanghai, China, p. 173

CLEMENTE, A., MARÍN-MANZANO, M. C., JIMÉNEZ, E. SONNANTE, G., MORENO, F. J., RUBIO, L. A. and DOMONEY, C. (2010) The cytotoxic effects of Bowman-Birk proteins from legumes on colon cancer cells are related to their ability to inhibit serine proteases. In: Legumes for global health, Legume crops and products for food, feed and environmental benefits. Abstract of the Fifth International Food Legumes Research Conference (IFLRC V) & Seventh European Conference on Grain Legumes (AEP VII), Antalya, Turkey, p. 22

CLEMENTE, A., MORENO, F.J., MARÍN-MANZANO, M.C., JIMÉNEZ, E. and DOMONEY, C. (2010) The cytotoxic effect of Bowman-Birk isoinhibitors**,** IBB1 and IBBD2, from soybean (*Glycine max*) on HT29 human colorectal cancer cells is related to their intrinsic ability to inhibit serine proteases. Mol. Nutr. Food Res. **54**, 396-405.

CLEMENTE, A., SONNANTE, G. and DOMONEY, C. (2011) Bowman-Birk inhibitors from legumes and human gastrointestinal health: current status and perspectives. Curr. Prot. Pep. Sci. **12**, 358-373

COTTAGE, A., GOSTKIEWICZ, K., THOMAS, J.E., BORROWS, R., TORRES, A.-M. and O'SULLIVAN, D.M. (2012) Heterozygosity and diversity analysis using mapped single nucelotide polymorphisms in a faba bean inbreeding programme. Mol. Breeding **30**, 1799–1809

CRUTZEN, P.J., MOSIER, A.R., SMITH, K.A. and WINIWARTER, W. (2008) N_2O release from agro-biofuel production negates global warming reduction by replacing fossil fuels. Atmos. Chem. Phys. **8**, 389–395. www.atmos-chem-phys.net/8/389/2008/

DOMONEY, C., CHINOY, C., PILLINGER, W., HASENKOPF, K., WILD, F., WARKENTIN, T., CLEMENTE, A. and CHARLTON, A. (2010) The genetic control of seed quality traits that impact on food and feed uses of pea. In: Legumes for global health, Legume crops and products for food, feed and

environmental benefits. Abstract (Keynote) of the Fifth International Food Legumes Research Conference (IFLRC V) & Seventh European Conference on Grain Legumes (AEP VII), Antalya, Turkey, p. 53

DOMONEY, C., KNOX, M., MOREAU, C., AMBROSE, M., PALMER, S., SMITH, P., CHRISTODOULOU, V., ISAAC, P.G., HEGARTY, M., BLACKMORE, T., SWAIN, M. AND ELLIS, N. (2013) Exploiting a fast neutron mutant genetic resource in *Pisum sativum* (pea) for functional genomics. Funct. Plant Biol. **40**, 1261-1270

FULTON, T.M., CHUNWONGSE, J., and TANKSLEY, S.D. (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants**.** Plant Molecular Biology Reporter **13**, 207-209

HELLENS, R.P., MOREAU, C., LIN-WANG, K., SCHWINN, K.E., THOMSON, S.J., FIERS, M.W.E.J., FREW, T.J., MURRAY, S.R., HOFER, J.M.I., JACOBS, J.M.E., DAVIES, K.M., ALLAN, A.C., BENDAHMANE, A., COYNE, C.J., TIMMERMAN-VAUGHAN, G.M. and ELLIS, T.H.N. (2010) Identification of Mendel's white flower character. PLoS ONE **5**, e13230

KHAMASSI, K., JEDDI, F.B., HOBBS, D., IROGOYEN, J., STODDARD, F., O'SULLIVAN, D.M. and JONES, H. (2013) A baseline study of vicine–convicine levels in faba bean (Vicia faba L.) germplasm, Plant Genetic Resources: Characterisation and Utilitisation. Plant Genet. Resourc. **11,** 250

KNOX, M., DOMONEY, C., COYNE, C., AMBROSE, M. and ELLIS, N. Relationship patterns in *Pisum* assessed with a range of molecular marker types (2010) In: Legumes for global health, Legume crops and products for food, feed and environmental benefits. Abstract of the Fifth International Food Legumes Research Conference (IFLRC V) & Seventh European Conference on Grain Legumes (AEP VII), Antalya, Turkey, p. 159, 2010

KNOX, M., MOREAU, C., LIPSCOMBE, J., DAVID BAKER, D. and ELLIS, N. (2009) High-throughput retrotransposon-based fluorescent markers: improved information content and allele discrimination. Plant Methods **5**:10 doi:10.1186/1746-4811-5-10

LORIEUX, M. (2012) MapDisto: fast and efficient computation of genetic linkage maps. Molecular Breeding **30**, 1231-1235

PANG, Y., WENGER, J.P., SAATHOFF, K., PEEL, G.J., WEN, J., HUHMAN, D., ALLEN, S.N., TANG, Y., CHENG, X., TADEGE, M., RATET, P., MYSORE, K.S., SUMNER, L.W., MARKS, M.D. AND DIXON, R.A. (2009) A WD40 repeat protein from Medicago truncatula is necessary for tissue-specific anthocyanin and proanthocyanidin biosynthesis but not for trichome development. Plant Physiol. **151**, 1114-1129

RUBIALES, D., AMBROSE, M. J., DOMONEY, C. and BURSTIN, J. (2011) Pea (*Pisum sativum* L.). In: Genetics, Genomics and Breeding of Cool Season Grain Legumes (Genetics, Genomics and Breeding in Crop Plants. Eds. M.P. de la Vega, A.M. Torres, J.I. Cubero and C. Kole). Science Pubs USA, New Hampshire, Jersey, Plymouth, pp. 1-49

SAINSBURY, F., THUENEMANN, E.C. and LOMONOSSOFF, G.P. (2009) pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. Plant Biotech. J. **7**, 682-693 doi: 10.1111/j.1467-7652.2009.00434.x

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 Physiol. **146**, 74-82

VOVLAS, N., TROCCOLI, A., PALOMARES-RIUS, J.E., DE LUCA, F., LIÉBANAS, G., LANDA, B.B., SUBBOTIN, S.A. and CASTILLO, P. (2011) *Ditylenchus gigas* n. sp. parasitizing broad bean: a new stem nematode singled out from the *Ditylenchus dipsaci* species complex using a polyphasic approach with molecular phylogeny. Plant Pathology **60:** 762–775

YOUNG, N.D. et al. (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature **480**, 520-524.

Publications in preparation include:

CLEMENTE, A., ARQUES, M.C., DALMAIS, M., LE SIGNOR, C., CHINOY, C., LAWSON, D.L.,

BENDAHMANE, A. and DOMONEY, C. (2014) Targetted induced mutations in a double-headed protease inhibitor gene in *Pisum sativum* L. reduce significantly the inhibition of target enzymes

WEBB, A. et al. (2014). A SNP based consensus genetic map for synteny-based trait targeting in faba bean (*Vicia faba*).

A list of others under development is available on request.

Summary tables referred to in text

Table 1 summarises the significant QTL and associated genetic traits, of which some key data are discussed above (p. 6 – 8).

Table 1: Summary of QTL detected for yield, plant and seed traits in pea

Trait	Population (allele with positive effect on trait)	Linkage group (LOD)	Near marker
Yield	Brutus x Enigma (E)	I (4.6 - 5.2)	Spds1
Yield	Kahuna x Enigma (E)	I (3.2)	AgpS2
Yield	Brutus x Enigma (E)	III (3.3 – 8.7)	LaDella / TT201
Yield	Kahuna x Brutus (B)	III (2.5 – 2.8)	LaDella
Yield	Kahuna x Enigma (E)	III (2.6)	TT201 / CA438/218
Yield	Kahuna x Enigma (E)	V (3.4)	AC358
Seed size	Kahuna x Brutus (K)	I (17 – 27)	AgpS2
Seed size	Kahuna x Enigma (K)	I (16 – 18)	AgpS2
Seed size	Brutus x Enigma (E)	I (5 – 7)	AT152/108-
Seed size	Kahuna x Brutus (K)	IV (3.2)	AT977
Seed size	Kahuna x Enigma (K)	IV (4 – 5)	AT975
Seed size	Kahuna x Enigma (K)	V (4 – 5)	GC327/386- / TA519
Seed size	Brutus x Enigma (B)	V (3 – 4)	TA519
Plant height	Kahuna x Brutus (B)	III (12.3)	LaDella
Plant height	Brutus x Enigma (B)	III (8 – 15)	LaDella
Plant height	Kahuna x Enigma (K)	III (6.0)	GC203/232+
Plant height	Kahuna x Brutus (B)	II (3.2)	AG235/183-
Plant height	Brutus x Enigma (B)	II (2.6)	AG235/183-
Plant height	Kahuna x Enigma (E)	II (3.9)	CA383
Plant height	Kahuna x Brutus (K)	IV (4.6)	*xyft*
Plant height	Kahuna x Brutus (K)	VI (4.2)	GA354/222+
Plant height	Kahuna x Enigma (E)	VI (2.9)	Susy3-1_Hha1
Standing	Kahuna x Brutus (*K*)	III (2.8)	AT513
Standing	Brutus x Enigma (*E*)	III (3.5)	GA807
Standing	Kahuna x Enigma (E)	III (3.3)	GC194/TT330/78+
Standing	Kahuna x Enigma (K)	V (3.4)	AC110
Standing	Kahuna x Brutus (*B*)	VI (2.9 – 3.3)	AG950/AT443
Standing	Brutus x Enigma (*B*)	VI (2.6)	TT440/95+/AT443
Standing	Kahuna x Enigma (E)	VI (2.9)	AA204
Standing	Kahuna x Brutus (*B*)	VII (2.8)	TC526

Linkage groups showing significant LOD scores determined by interval mapping are given, where the LOD thresholds of significance were in the range 2.4 – 2.6. The nearest marker is listed for every QTL. (For standing ability, various measures were made, including lodging scores where high lodging equates to poor standing ability. For alleles associated positively with lodging, the alternative allele is listed above (italicised) for positive standing).

Table 2 summarises the PCGIN expanded genetic resources for pea and faba bean, some of which are discussed above (p. 11 - 14).

Table 2: Summary of expanded genetic resources available for pea and faba bean

Species	Resource	Implication	Reference
Pea	PA2 variant	Improved seed quality	Vigeolas et al. 2008
Pea	TI1 (TILLING)	Improved seed quality	Clemente et al. (in prep)
Pea	FN deletions	Altered seed composition	Domoney et al. 2013
Pea	Sgr variants (TILLING)	Improved seed colour stability	In preparation
Pea	Sgr variants (germplasm)	Improved seed colour stability	In preparation
Pea	DM resistant line	Downy mildew resistance	-
Pea	Kahuna x Enigma RILs	Mapping yield traits	-
Pea	Brutus x Kahuna RILs	Mapping yield traits	-
Pea	Brutus x Enigma RILs	Mapping yield traits	-
Pea	Princess x 185 RILs	Mapping seed traits	Chinoy et al. (in prep)
Faba bean	900 accessions	Resource for genetic variation	-
Faba bean	130 SSD lines	Pure bred lines	-
Faba bean	BPL10 x Albus	Mapping population (general)	-
Faba bean	29H x Albus	Mapping population (general)	-
Faba bean	NV643 x NV648	Mapping resistance to *Ditylenchus gigas*	(Vovlas et al. 2011)[1]
Faba bean	NV657 x NV643	Mapping resistance to *Ditylenchus gigas*	(Vovlas et al. 2011)[1]
Faba bean	NV639 x NV658	Mapping population (general)	-

[1] *reference for information on Ditylenchus gigas*