# Implementation of Markers for Pulses (iMAP)

**Dr. Bunyamin Tar'an**
University of Saskatchewan – CDC; National Research Council of Canada – PBI

| SPG Contributions | Project Status | Duration/Timeline of Project (Year to Year) | Total Project Cost |
|---|---|---|---|
| $2,136,008.00 | Completed | August 2010 – August 2013 | $2,136,008.00 |

## Project Description

Saskatchewan contributes 97% of Canada's lentil crop, 99% of the chickpea crop and 72% of the Canada's field pea crop. Continued expansion of the pulse crop acreage in Canada is expected, combined with more intense rotations. This requires the crops to have a wider range of adaptation while maintaining high yield, diverse product range, and acceptable quality. The expansion has also increased the risk from foliar and soil borne pathogens, posing a real threat to the sustainability of crop production. New techniques, which augment conventional breeding practices, are needed to facilitate the rapid incorporation of traits into new cultivars that will allow growers to improve their production systems, increase the stability of the crops, and move commercial yield forward.

The advent of genomic technologies has helped breeders and geneticists to understand complex traits, identify genomic regions controlling traits of economic and agronomic importance, and design strategies for improvement of these traits with greater precision and efficiency by marker assisted selection. The application of these molecular breeding technologies requires genomic resources that allow detailed characterization of available genotypes.

The iMAP project applies the latest genetic tools to enhance the current breeding program by marker-assisted selection. The genomic resources will be used to establish a set of molecular tools and integrate them with genetic materials from the CDC breeding program. This will provide greater precision in selection and lead to faster development of improved cultivars.

This study used a related series of strategies to identify and map large numbers of molecular markers, for lentils, peas, chickpeas, common beans, and its close relative, tepary beans.

Initially, a SNP genotyping platform was developed for each crop, with the purpose of identifying and mapping molecular markers and relating them to the map of the well characterized legume, Medicago truncatula. SNPs have become the marker of choice due to their abundance in plant genomes.

The Illumina GoldenGate SNP genotyping platform is a highly multiplexed ligation based SNP assay. Up to 1,536 loci can be genotyped simultaneously. This provides a very robust platform for global SNP genotyping in a diploid genome and thus it is an ideal format for rapid genetic mapping of many SNP loci in the pulse crop genomes.

SNPs were identified using a variety of materials for each crop, and this information used to design a unique 1536 locus GoldenGate assay for each crop (for bean, 768 from each of common and tepary bean were used). Genotypes were tested and were used with other relevant genetic information to generate maps and linkage groups, which could then be used to analyze parental lines of different inbred populations of the crop and analyze the diversity of the sample panels for each crop.

At the same time, core sets of genotypes for each species were placed in three-year multi-location replicated field trials to provide robust phenotypic data for genetic analysis of economic traits by means of quantitative trait loci (QTL) analysis and association mapping. The QTL could be aligned with the SNP data and converted into 'breeder friendly' markers as KASP (Kompetitive Alelle Specific PCR) assays, which can be run with simpler equipment and used for routine breeding. Its use could be extended into rapid generation advancement using tissue culture, which was tested using the example of disease resistance in bean.

The data and resources were placed in the public domain, including Knowpulse (knowpulse.usask.ca), our own portal for all genetic and genomic information database.

## Outcome

For each species, a reference genotype was chosen and several others compared to it, generating large series of SNPs.

For lentils, the reference genome was CDC Redberry and 10 others were compared to it. A subset of SNPs was selected according to several criteria, and used to design the 1536 Golden Gate assay. It was then used to genotype a population derived from a cross between CDC Robin and 964a-46, two lines with diverse genetic backgrounds: one is a small red lentil and the other a large green lentil, and the population segregates for multiple traits. The data (484 SNPs) were mapped into seven linkage groups corresponding to the seven chromosomes of lentil. The linkage groups range from 58 cM to 226 cM, and together they cover 834.7 cM. This map, the most comprehensive lentil map to date, will form the basis for further genome assembly derived from the on-going genome sequencing initiative in lentils.

For peas, a reference genome (CDC Bronco) and seven others were compared, and available information from public sequence data for this and other legume species was added, to generate a large set of potential SNPs and select those for the 1536 GoldenGate assay. Five Recombinant Inbred Lines (RIL) populations were then examined with the assay, and a high level of polymorphism was found. The data for 945 loci were mapped into seven linkage groups corresponding to the seven pea chromosomes covering 772.0 cM of the consensus map.

For chickpeas, a similar strategy with CDC Frontier was used to provide 1536 GoldenGate assay SNPs. A total of 92 RILs from a cross between ICCV96029 x CDC Frontier were used to construct the linkage map. A total of 1,336 SNPs was mapped into eight linkage groups corresponding to the chickpea chromosomes, covering a genetic distance of 652.9 cM. A draft genome of Kabuli and Desi genotypes was available, which allowed comparisons between physical and map distance, and estimates of recombination rate in different parts of the genome. Some regions were found to be "hotspots" with elevated recombination rates.

For beans, comparison of eight domesticated bean genotypes was used to develop the 768 GoldenGate assay SNPs, together with other genetic information. The bean assay was used to analyse RIL population derived from a cross of Expresso x Black Diamond, which led to 11 linkage groups corresponding to chromosomes, with a

total length of 437.4 cM.

Tepary bean information was used to contribute to the bean genotyping effort. Eight domesticated or semi-domesticated genotypes were used to develop the 768 GoldenGate assay SNPs. The tepary bean map covered a genetic distance of 1103 cM arranged in 11 linkage groups. The tepary bean and common bean have a high degree of collinearity, as would be expected from their near relationship.

A considerable, but lesser, degree of collinearity is also found among all the pulse genomes. When compared to the Medicago genome and among themselves, long regions of collinearity are found, with some major changes in arrangement such as translocations and inversions. These regions of collinearity are important in that they allow genetic information to be inferred among genomes. A critical gene or QTL located in a gap in one genome may often be identified using information derived from another pulse species.

Phenotypic assessment of core mapping populations and panel genotypes for association mapping was done across multiple locations and years. For each species, a set of 94-147 RILs from a cross between moderately different lines was grown in replicated trials over two or three years, with two locations, and detailed phenotypic data was collected. Some of the traits analyzed included growth form and maturity times, standability, lodging, seed size, shape, colour and mineral content, yield, and in some cases, disease response. Within each RIL population, association mapping could be carried out for these traits. Genetic distances for different traits could be found and QTL identified. This data was then used to align physical and genetic maps, and find the nearest SNP markers to the candidate QTL region. For each species, a large panel of lines was then examined by association mapping. Where pre existing groups were known, such as market class in lentils, there tended to be a strong correlation. QTLs were identified which corresponded to, for example, higher iron or zinc content, and related SNPs could be identified. A set of high value SNPs linked to key agronomic and quality traits was identified for each species.

The GoldenGate SNP markers associated with QTLs were converted into 'breederfriendly' markers as KASP assay, which allows rapid genotyping of targeted subsets of SNP markers in breeding populations. These KASP markers consisted of two groups: first, those SNPs that are associated with the QTL of the important traits; and second, sets of SNPs equally spaced along each linkage group of the map, for use in mapping future populations. In total, 1,828 KASP markers were designed across lentils, peas, chickpeas, common beans, and tepary beans. After validation, the KASP markers associated with key agronomic traits were used to genotype parents and F1 (and backcross) and F2 progeny, as proof of concept of MAS prior to implementation of marker technology in routine breeding.

The information from QTL and marker associations can be used to identify if markers linked to particular genes or QTLs are present in a proposed parental line, in the F2, in the F1 of a multi-parent cross, or in any other convenient genotype. DNA from all members of a segregating population can be quickly prepared and tested using KASP assays, and the early-generation plants with the trait of interest can be identified and others discarded, long before the trait can be detected in the whole plant or its seed. This allows much greater efficiency in plant breeding, by keeping only the progeny which are most likely to have the trait in question. In the case of disease resistance, it reduces the use of expensive and somewhat unreliable disease nurseries.

For example, high value SNPs and their use in MAS were used to screen for resistance to common bacterial blight and anthracnose in common beans. SNPs were identified which were closely linked to disease resistance genes, and plants carrying the resistance genes could be identified in the segregating populations.

The markers were similarly tested in lentils, in a search for parents to use for transmission of herbicide (imidazolinone) tolerance, and to identify backcross progeny with the trait. The scores in both cases had a good match with the cultivar's or progeny's reaction to the herbicide. A comparable trial was also done in chickpeas.

The integration of the KASP assay with a tissue culture system for rapid generation advancement for selection of disease resistance in common beans has been explored. This permits plants in tissue culture to be advanced for several generations without exposure to the disease.

The use of the markers will be extended further to more quantitative traits or those with less clear-cut phenotypes as more data become available from controlled trials.

The developed data base for 454 transcriptome sequencing, generated EST database, and SNP resources for all four crops has been made available, and the results are being published, in open access international journals. All data and resources developed are in the public domain and are available at KnowPulse (http://knowpulse2.usask.ca/portal).

The research has successfully completed the development of core linkage map and QTL analysis for each pulse crop and genotyping the association mapping panels of each species using the 1,536 SNP GoldenGate array for each of lentil, pea, chickpea and 768 SNPs for each of common bean and tepary bean. The markers identified in this project will be used to advance plant breeding of these pulse species more efficiently and with higher precision.

The total outputs of the iMAP project help us to reach the long term goal of increasing the precision with which breeding populations are evaluated, which will facilitate the rapid development of improved pulse cultivars.

## Research Objective

**OBJECTIVE 1**
To use the genomic resources to establish a set of molecular tools and integrate them with genetic materials from the Crop Development Centre (CDC) breeding program to provide greater precision in selection.