AAFC RESEARCH BRANCH
Research Project Final Report

Developing Innovative Agri-Products Program (Vote 10 Funding)

Project Title:	Activity A.2: Develop techniques that will help simplify selection breeding of large populations
Start Date (yyyy-mm-dd):	2010-04-01
Expected End Date (yyyy-mm-dd):	2013-01-31
Actual End Date (yyyy-mm-dd):	2013-01-31
Principal Investigator (PI):	Neil Emery

Short Executive Summary of report:

I. Development of markers for cytokinin based yields in soybean

1. Identified key cytokinins (CKs) related to improvement of natto beans flowering, pod setting, and seed development through comparison of cytokinin biochemistry in natto soybean lines with varying levels of yields. This includes LC-MS/MS analysis of the cytokinin metabolites at flowering, pod setting and seed development stages.

2. Identified cytokinin genes and discovered potential molecular markers for increased yields in soybeans and natto beans through gene cloning and sequencing, gene expression analysis of cytokinin biosynthetic and catabolic genes.

These activities will lead to higher soybean yields through cytokinin gene marker-directed breeding for regular and natto soybeans.

II. Use of molecular markers for soybean improvement

Developed and validated molecular markers for high alpha-tocopherol contents, lipoxygenase free, SCN resistance, and different fatty acid compositions based on published publically available literature. These markers are now available for Sevita to apply to their breeding program.

A. Research Progress and Accomplishments (to date in relation to expected milestones and deliverables / outputs)

- Include brief summary of:
  - Introduction, literature review, objectives, milestones and deliverables / outputs.
  - Approach / methodology (summary by objectives).
- Include results and discussion (overview by objectives and milestones), next steps and references.

1. Summary of introduction, literature review, objectives, milestones and deliverables / outputs

This research provides the potential selection markers for greater soybean and natto bean yields through the identification of key genes controlling a major plant growth hormone group, the cytokinins (CKs). The research combines comprehensive cytokinin metabolite profiling and gene expression analysis to identify key cytokinins and cytokinin genes involved in reproductive development of soybean and natto beans. Potential molecular markers discovered through gene-specific polymerase chain reaction (PCR) and sequencing of multiple soybean and natto bean lines were validated and applied for soybean and natto bean breeding programs. The application of these plant hormone gene markers could lead to higher

yields, therefore, more efficient land resource use and growth in the domestic and international agri-food markets.

The impact of CKs in seed growth appears to be positive among species or families of plants tested to date (Ashikari 2005, Zalewski 2010, Powell 2013); but the extent of this impact remains unexplored for most species. Early work has shown that CKs accumulate during critical periods of cell division following anthesis in cereals (Morris 1997) and similar dynamics, are seen in legumes (Emery et al. 1998, Emery et al. 2000).

Soybean, in particular, is a promising target for CK-based yield improvements. Its fruit set was repeatedly shown to increase in response to CK applications (Nagel et al. 2001, Liu et al. 2004). In addition, the preliminary results with soybean indicate that its CK profile is predominantly comprised of cis-isomers, which are thought to have lower activities than corresponding trans-CKs. Therefore, the strategic increase in CK activity during seed development, either through increased concentration or by increased emphasis on trans-CK isomer production are plausible ways of achieving higher yields.

Ashikari et. al. (2005) discovered a quantitative trait locus (QTL) accounting for 40% increases in rice yields. This high yield QTL mainly involved a deficient enzyme - cytokinin oxidase (CKX) – that is responsible for the degradation of cytokinins. The Emery lab's work at Trent University with legumes showed that CKs accumulate and promote embryo growth *in vitro* (Emery & Atkins 2006, Quesnelle and Emery 2007). This indicates that CK related growth potential could match that of rice documented by Ashikari et al. (2005).

Zalewski et al. (2010) reported that silencing of the HvCKX1 gene resulted in increased accumulation of CKs in developing grains of barley and therefore increased barley grain weight and yields. Similarly, mutants of *Arabidopsis* with reduced expression of AtCKX3 and AtCKX5 genes showed higher number of inflorescences and dramatically increased seed productivity (Bartrina et al., 2011).

This research project was proposed to explore the potential of greater soybean and natto bean yields through the identification of key CKs and genes that control CK metabolisms, and develop CK gene markers for soybean and natto bean breeders to achieve yield improvement. The investigation combines comprehensive analyses of seed CK metabolites, gene expression, gene-specific PCR, and sequencing to identify yield-based markers in soybean and natto beans. These activities will lead to higher yields through cytokinin gene-marker directed breeding for soybeans and natto beans.

In addition, some molecular markers were developed and validated for several important agronomic traits such as high alpha-tocopherol contents, lipoxygenase free, soybean cyst nematode (SCN) resistance, and different fatty acid compositions based on publically available literature to accelerate soybean selection within Sevita's breeding programs.

Performance indicator:

A2.1 Development of markers for cytokinin based yields in soybean

Identification of key cytokinins related growth signalling genes.
Soybean lines with greater yields.
Identify DNA markers for breeders to direct their efforts towards increased yields.

A2.2: Development of molecular marker to assist in molecular selection of breeding lines

Milestones:

CK metabolite profiles (up to 20 forms of CK identified and quantified) were generated for the two earliest reproductive stages R1 and R2.
CK metabolite profiles were generated for reproductive development stages R3 and R4.
Characterization of cytokinin biosynthesis genes (*Glycine max* isopentenyl transferase (GmIPT)).
GmIPT expression quantified by quantitative real-time PCR (qRT-PCR).
Identify and test potential markers for high GmIPT activity.
Find and characterize cytokinin degradation genes (*Glycine max* cytokinin oxidase (GmCKX)).
Quantify the gene expression GmCKX family members by qRT-PCR.
Indentify and test potential DNA markers based on variation in sequences between low and high activity CKX enzymes.

Deliverables / outputs:

Month 1-6: LC-MS/MS analysis of the CK metabolites at development stages R1 and R2.
Month 7-10: LC-MS/MS analysis of the CK metabolites at development stages R3 and R4.
Month 11-13: Primer design, optimization and organ specific testing of expression.
Month 14-18: Analysis of IPT gene expression at each of the 4 stages of seed development by RT-Q-PCR.
Month 19-21: Sequence high activity seed IPT genes & test markers on seed library.
Month 22-25: Degenerate primer design, analysis of soybean homologs, specific primer design.
Month 26-29: Analysis of CKX gene expression at 4 stages of seed development by RT-Q-PCR.
Month 30-32: Sequence low activity seed CKX genes & screen seed library with markers.

2. Approach/Methodology (summary by objectives)

The Agilent HPLC and AB SCIEX QTRAP 5500 LC/MS/MS mass spectrometry was applied to identify key cytokinins (CKs) related to improvement of soybean and natto bean flowering, pod setting, and seed development through comparison of cytokinin biochemistry in soybean and natto bean lines with varying levels of yields. Harvested samples were kept at -80ºC before cytokinin extraction. The frozen tissues were ground and homogenized in cold (–20ºC) Bieleski extraction buffer before solid phase extraction procedure. Solid phase extraction was conducted with the use of Oasis MCX columns (60µm 6cc, 150mg) to separate different forms of CKs, and the CK nucleotide fraction was further purified by C18 columns (500mg, ODS) solid phase extraction. In the first single LC/MS/MS run, a total of 20 different cytokinins were analyzed (CK free bases, iP, c-Z, t-Z, DZ and their ribosyl, glucosyl, O-glucosyl and methylthio-conjugates). In the second LC/MS/MS run, four nucleotide forms were analyzed as ribosides following digestion with phosphatase. The 2H-labelled CKs were used as internal standards. The CK metabolite profiles were compared with seed size and yield data collected based on field trials in multiple years and with multiple replicates.

Whole genome blast search was conducted to identify the key cytokinin genes in soybean genome using *Arabidopsis* IPT (CK biosynthesis), CKX (CK degradation), and LOG (“lonely guy” gene from *Oryza sativa*; CK activation) genes as seed sequences. Gene-specific primers used for Real-time PCR were 20-22 nt long, GC content for each primer is between 50% and 55%, melting temperature (Tm) is 55-60^{o}C, and the size of the amplicon ranged from 90-100bp. When the gene size was over 3kb, a long range PCR was conducted to amplify the genomic DNA of these CK genes. Purified PCR products were sequenced using Sanger sequencing method and ABI 3730 DNA Analyzer. Sequencing results were analysed using MEGA 5 software and single nucleotide polymorphisms (SNPs) were verified by comparison the original chromatograms among soybean and natto bean lines. Entire sequences were analyzed for all members of the IPT, CKX, and LOG gene families to look for changes in coding sequence that may influence protein activity. The potential SNPs that caused amino acid change were analyzed for their potential severity of amino acid substitutions and protein 3D structure stability using

SWISS Model Workplace, SDM, and PyMOL programs.

Markers for high alpha-tocopherol contents, lipoxygenase free seeds, SCN resistance and different fatty acid compositions were developed and validated based on published publically available literature. These validated markers were applied for parental and progeny screening in the breeding programs of Sevita International.

3. Summary of results and discussion (overview by objectives and milestones)

I. Development of markers for cytokinin based yields in soybean and natto beans

Part 1. Cytokinin extraction, profiling, and data analysis

1. Plant tissue samples were collected at critical yield determining stages during the early reproductive growth periods R1-R5 from 20 natto lines.

   Twenty natto lines were used for the plant tissue collection through R1 to R5 (table 1). Four replicates were prepared at each stage for each natto line. Of these 20 natto beans, nine are high yield natto beans and 11 are low yield natto beans. They were grown at the AAFC Experimental Farm in Ottawa in the growing seasons of 2011 and some were re-harvested and analyzed in the growing seasons of 2012.

Table 1. Natto lines and cultivars used for cytokinin profiling

Name	Yield (kg/ha)	(Bu/ac)	Rank	Days to maturity
OT08-13	4030	60.5	5	132
X5076-1-1-1-209-B	3750	56.2	6	126
X5076-1-1-1-35-B	3679	55.2	7	126
X5076-1-1-1-133-B	3649	54.7	8	122
Apalis	3605	54.0	9	133
X5076-1-1-1-159-B	3549	53.2	10	126
X5076-1-1-1-121-B	3522	52.8	11	130
Chikala	3445	51.7	12	122
X5076-1-1-1-108-B	3424	51.4	13	123
AC Colibri	2526	37.9	138	110
X5076-1-1-1-277-B	2516	37.7	141	124
X5076-1-1-1-6-B	2515	37.7	142	114
X5076-1-1-1-131-B	2471	37.0	143	121
X5076-1-1-1-68-B	2443	36.7	144	122
X5076-1-1-1-29-B	2380	35.7	145	114
X5076-1-1-1-283-B	2354	35.3	146	108
X5076-1-1-1-81-B	2329	34.9	147	124
X5076-1-1-1-95-B	2256	33.9	148	113
X5076-1-1-1-239-B	2216	33.2	149	109
X5076-1-1-1-199-B	2157	32.4	150	121

2. Cytokinins were extracted and purified from all the early reproductive growth period R1-R5 samples.
   Harvested samples were kept at -80ºC before cytokinin extraction. The frozen tissues were ground and homogenized in cold (–20ºC) Bieleski extraction buffer before solid phase extraction procedure.
   Solid phase extraction was conducted with the use of Oasis MCX columns (60µm 6cc, 150mg) to separate different forms of CKs, and the CK nucleotide fraction was further purified by C18 columns (500mg, ODS) solid phase extraction.

3. Cytokinin profiling was finished for the early reproductive growth period R1-R5 samples.

   The CK fractions extracted were analyzed using the Agilent HPLC and AB SCIEX QTRAP 5500 LC/MS/MS system.

4. Data analysis of cytokinin profiling was completed for the early reproductive growth period R1-R5 samples.

   Multiple types of cytokinins including isopentenyl adenine (iP), trans-Zeatin (transZ), cis-Zeatin (cisZ), dehydro-Zeatin (DHZ) and their precursors were detected in early reproductive tissues through R1-R5 stages. Trans-Zeatin (transZ) is considered the most active CK and it was the dominant type of cytokinin in young pods of natto beans (figure 1).

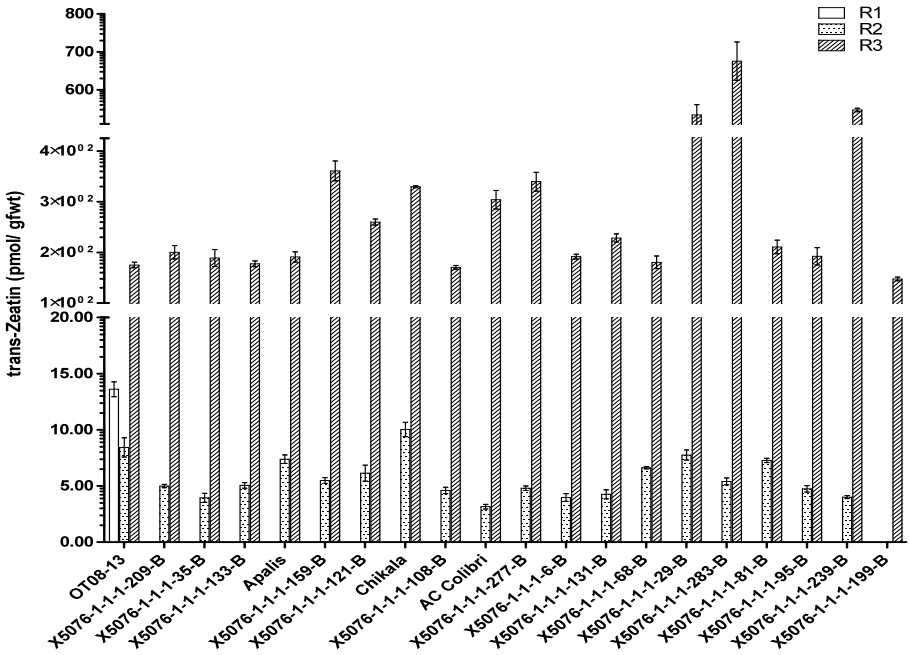

Figure 1. Trans-Zeatin (the most active CK) was identified as the dominant cytokinin during reproductive stages

Part 2. Cytokinin gene discoveries in soybean genome, cloning, expression profiling, sequencing, and potential markers

1. Genes involved in cytokinin biosynthesis, oxidation (degradation) and activation were discovered in the soybean genome by whole genome blast search in soybean using *Arabidopsis* IPT genes, rice CKX genes, and rice LOG genes as seed sequences.

A. Genes involved in CK biosynthesis and gene phylogenetic analysis

A total of 14 CK biosynthesis genes (Gm-IPT) were found in soybean genome. Phylogenetic analysis results show that Gm-IPT genes clustered in six groups (figure 2).

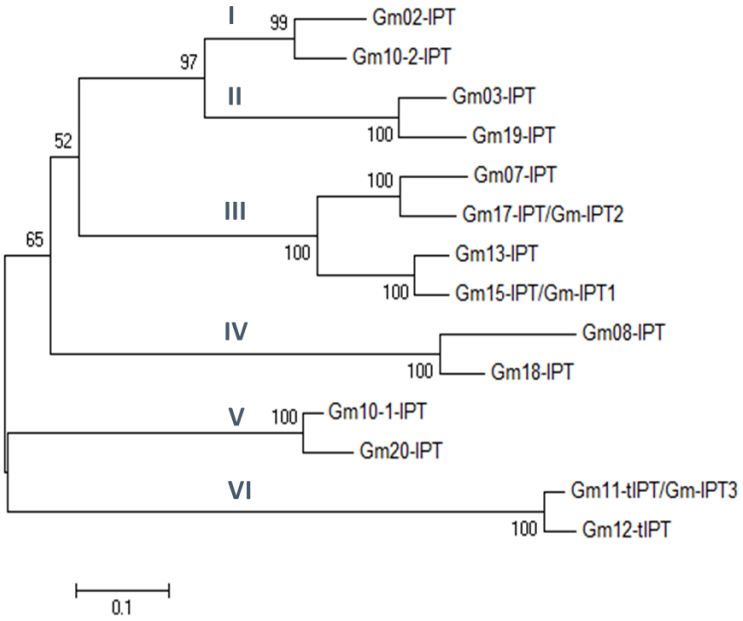

Figure 2. Phylogenetic analysis of soybean cytokinin biosynthesis genes (Gm-IPT)

B. Genes involved in CK degradation and gene phylogenetic analysis

A total of 17 CK degradation genes (Gm-CKX) were found in soybean genome. Phylogenetic analysis results show that Gm-CKX genes also clustered in six groups (figure 3).

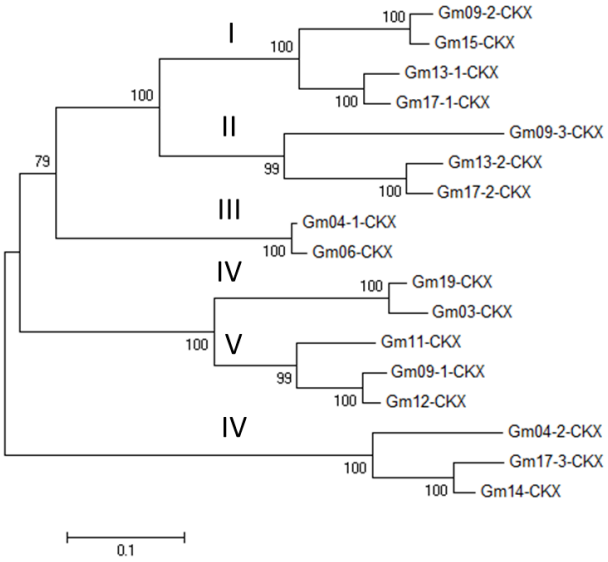

Figure 3. Phylogenetic analysis of soybean cytokinin degradation genes (Gm-CKX)

C. Genes involved in CK activation and gene phylogenetic analysis

A total of 22 CK activation genes (Gm-LOG) were found in soybean genome. Phylogenetic analysis results show that Gm-LOG genes clustered in seven groups (figure 4).

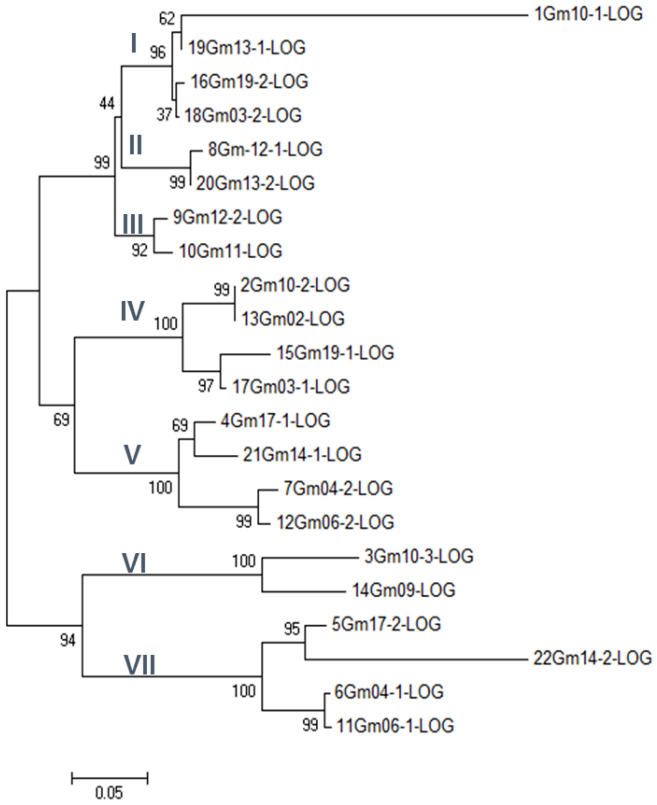

Figure 4. Phylogenetic analysis of soybean cytokinin activation genes (Gm-LOG)

2. Identification of key cytokinin genes in young seeds through gene expression analysis

Primers for 14 Gm-IPT and 16 Gm-CKX genes' expression analyses were designed, synthesized and optimized. Primers used for Real-time PCR were designed using Eurofins' Primer Design Tools. Primer sequences were 20-22 nt long, GC content for each primer was between 50% and 55%, melting temperature (Tm) was 55 – 60◦C, and the size of the amplicon ranged from 90-100bp.

Gene expression analysis was conducted for all these 14 Gm-IPT and 16 Gm-CKX genes in soybeans. Gene expression analysis of 14 Gm-IPT and 16 Gm-CKX genes were completed by using high quality RNA from seeds of three developmental stages (R5, R5.5, and R6) of regular soybean; High and low activity Gm-IPT genes were identified.

A. Gm-IPT gene expression analysis in soybean young seeds (R5, R5.5, and R6 stages).

Out of 14 Gm-IPT genes, six Gm-IPT genes have relatively high expression levels at three seeding stages. Gm08-IPT and Gm10-2-IPT genes have relative high and stable expression at different seed development stages; Gm18-IPT and Gm-tIPT3 (a tRNA IPT) have higher

expression levels in young seeds (R5) than developed seeds (R5.5 and R6 stages); Gm-IPT2 showed declining expression at R6 stage; Gm12-tIPT, another tRNA IPT gene, dramatically increased expression levels during seed development from R5 to R5.5 to R6 stage (figure 5).

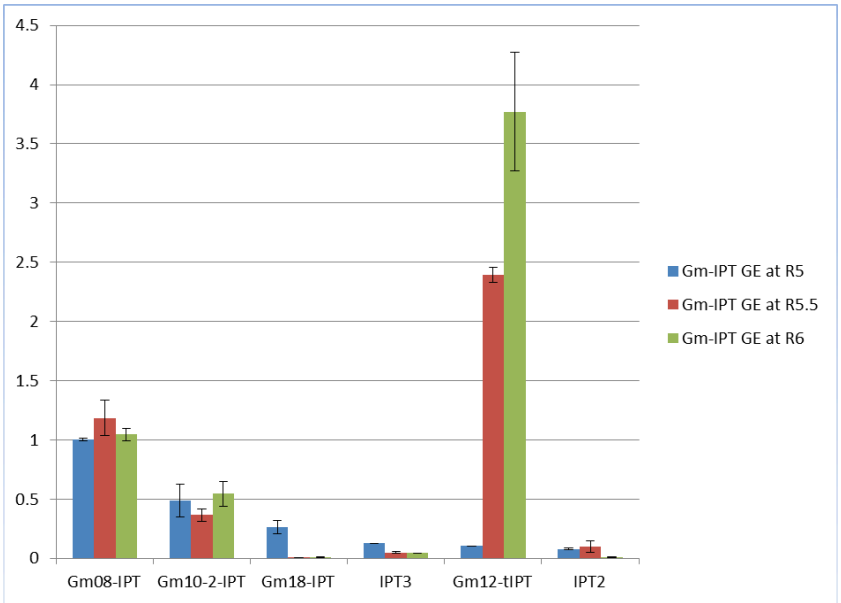

Figure 5. The gene expression level change of six Gm-IPT genes at three seed developmental stages

2b. Gm-CKX gene expression analysis was completed in soybean young seeds (R5, R5.5, and R6 stages).

Out of 16 Gm-CKX genes, six of them have relatively high expression levels in soybean seeds ((R5, R5.5, and R6 stages): Gm14-CKX, Gm03-CKX, Gm19-CKX, Gm04-1-CKX, Gm04-2-CKX, and Gm15-CKX. Gm14-CKX and Gm19-CKX showed increased expression levels along with the seed development from R5 to R6 stages; Gm14-CKX has the highest expression levels in all of the seeding stages than other Gm-CKX genes (figure 6).

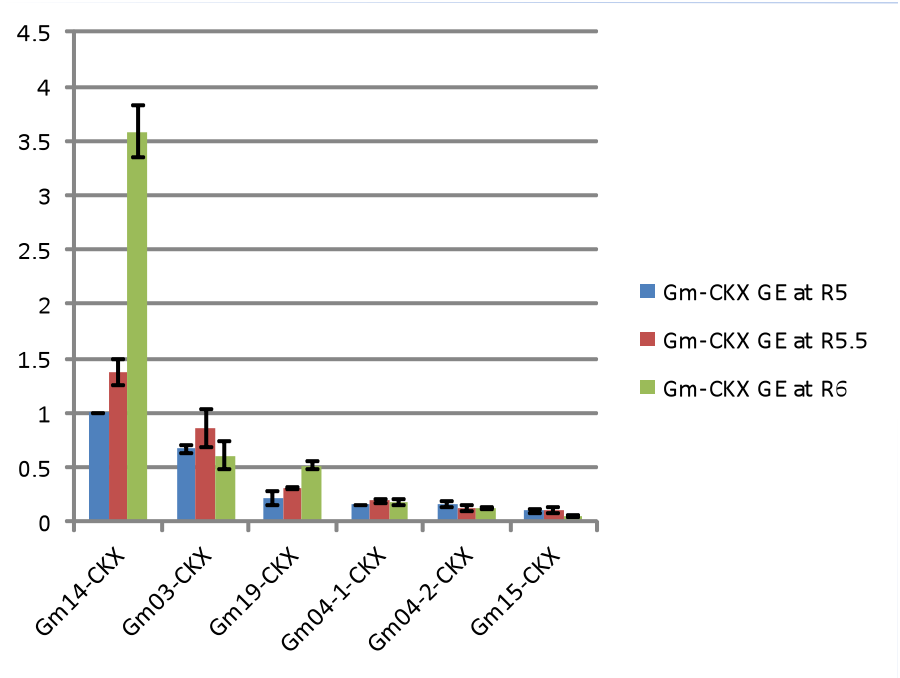

Figure 6. The gene expression level change of 6 Gm-CKX genes at three seed developmental stages

3. Enough high quality genomic DNA for cytokinin genes' cloning was extracted from 20 soybean lines including seven natto lines. Young leaves of soybeans and natto beans were harvested and the high quality DNA was extracted using the E.Z.N.A. Plant DNA Kit or QIAGEN DNeasy Plant Mini Kit. Primers were designed by using the Primer 3 Plus program with some optimizations based on the Gm-IPT, CKX, and LOG gene's genomic sequences. The PCR were conducted to clone cytokinin genes. The QIAGEN Quick PCR Purification kits were used for PCR products purification. The quality and quantity of purified PCR products were detected by both electrophoresis and Nanodrop spectrophotometer. Purified PCR products of cytokinin genes were sequenced using the BigDye Terminator v3.1Cycle Sequencing Kit and the ABI 3730 DNA analyzer.

   The following soybean and natto bean lines and cultivars were used for the cytokinin gene marker project:
   PI 88788
   OAC Wallace
   OAC Bayfield
   Keszthelyi A.S.
   RCAT Ruthven
   Hefeng 25
   IA1010LF
   DH410
   OAC 06-14
   DH420
   DH3604
   Kinusayaka
   OT08-13
   X5076-159B

Chikala
X5076-108B
AC Colibri
X5076-68B
DH748
OAC Calypso

4. Primers were designed for cloning and sequencing of 14 Gm-IPT genes, 17 Gm-CKX genes, and 22 Gm-LOG genes. A total of 14 Gm-IPT genes, 17 Gm-IPT genes and 22 Gm-LOG genes were cloned and purified from 20 soybean and natto lines.
5. Completed sequencing of 14 Gm-IPT genes cloned from 12 soybean lines including a natto line and a small seed landrace line, completed sequencing of 17 Gm-CKX genes cloned from 20 soybean lines including seven natto lines and completed sequencing of 16 Gm-LOG genes cloned from 20 soybean lines including seven natto lines.
6. Four SNPs with amino acid changes were discovered among 14 Gm-IPT genes. A yield-informative SNP marker was developed for Gm12-tIPT gene.

Table 2. Four SNPs with amino acid changes found in Gm-IPT genes

SNP	Gene Name	DNA polymorphism	Amino acid change	Note	Marker Developed
1	Gm18-IPT	T → C, 419 bp from ATG	Leucine → Proline, 140 aa from Methionine	The mutated residue is located in a domain that is important for binding of other molecules. Mutation of the residue might disturb this function.	
2	Gm19-IPT	G → T, 509 bp from ATG	Glutamic acid → Aspartic acid, 169 aa from Methionine	Might not be a significant change	
3	Gm12-tIPT	C → T, 188 bp from ATG	Proline → Leucine, 63 aa from Methionine	Mutant residue is located near a highly conserved position. The mutation might abolish the required flexibility of the protein at this position. The SNP has strong association with yield.	Yes
4	Gm12-tIPT	T → C, 724 bp from ATG	Serine → Proline, 242 aa from Methionine	Might not be a significant change	

7. Total of 10 SNPs with amino acid changes were discovered in 17 Gm-CKX genes. Further characterizations are ongoing, and new markers will be developed for selection.
8. Sequencing of 16 LOG genes is completed and no SNPs with amino acid changes were found yet.

Table 3. Ten SNPs with amino acid changes found in 17 Gm-CKX genes, markers under development

SNP	Gene name	Polymorphism	Position	Nucleotide change on CDS	Notes	Varieties
1.	GmCKX09-1	Leucine - Phenylalanine	330	A-T, 990	Both Hydrophobic, aliphatic to aromatic amino acid	DH3604, OT 08 13 and X5076 159B
2.	GmCKX09-1	Leucine - Glutamine -	509	T-A, 1526	Has an extra oxygen to deal with	DH3604, OT 08 13 and X5076 159B
3.	GmCKX06	Histidine - Arginine	16	A-G, 35	Not conserved region	PI88788, X5076 68B and X5076 159B
4.	GmCKX06	Lysine - Glutamic acid	77	A-G, 217	Highly basic to acidic. May cause destabilization.	PI88788, X5076 68B and X5076 159B
5.	GmCKX04-1	Threonine - Proline	419	A-C, 1282	Aliphatic to aromatic. Solvent accessibility decreased. Stabilizing mutation. Expression levels may increase.	DH3604, X5076 108B, Keszthelys AS, OAC 06-14, PI88788
6.	GmCKX17-1	Proline - Serine	37	C-T, 109	Not conserved region	PI88788
7.	GmCKX04-2	Alanine-Valine	66	C-T, 194	Uncertain! SDM prediction says it's a destabilizing mutation. Protein does not have a GHS region.	DH3604, X5076-108B
8.	GmCKX14	Threonine - Alanine	69	A-G, 205		OAC Wallace, K.A.S., IA1010LF, DH420, DH748
9.	GmCKX14	Histidine - Glutamine	105	C-G, 315		OAC Wallace, K.A.S., IA1010LF, DH420, DH748
10.	GmCKX14	Alanine -Glycine	159	C-G, 476	Center of the protein. Slightly destabilizing and may cause substrate affinity issues.	OAC Wallace, K.A.S., IA1010LF, DH420, DH748

II. Use of molecular markers for soybean improvement

1. Validation and application of molecular markers for soybean seed lipoxygenase free screening

   Soybean seed lipoxygenase gene-specific InDel (insertion/deletion) markers and SNP (single nucleotide polymorphism) markers were validated and applied for both parental screening and progeny identification and selection. These markers have 100% accuracy for lipoxygenase free

test results.

2. Development, validation, and application of molecular markers for SCN resistance screening
   A. A co-dominant SNP marker for *rhg1* (resistant to *Heterodera glycines*) locus SCN resistance selection

      The first SCN marker applied to Sevita parental lines is a co-dominant SNP marker developed based on *rhg1* locus SNP haplotypes, that could differentiate these major SCN resistance sources such as PI 548402 (Peking), PI 88788 (DH410), and PI 437654 from a group of susceptible soybean lines such as OAC Wallace, OAC Bayfield, IA1010LF, Hefeng 25, OAC 06-14, DH420, and DH3604 et al.

   B. multiplexing marker for *rhg1* locus SCN resistance selection combined with internal control gene

      A new multiplexing marker was validated for SCN resistance screening based on the discoveries of three genes with multiple repeats for SCN resistance within the *rhg1* region. An endogenous gene as internal control is used for positive PCR detection considering that some samples have poor DNA qualities e.g. when deriving from old leaves or seeds.

3. Development of molecular markers for tocopherol improvement in soybean seeds
   A. 35 polymorphic SSR markers tightly linked to soybean tocopherol genes and six tocopherol gene-specific markers were discovered. These markers will be a useful molecular marker pool for high tocopherol soybean development.
   B. An alpha tocopherol gene specific marker was developed based on the promoter sequence differences of gamma-TMT3 gene, which encodes gamma-tocopherol methyltransferase, the gene responsible for high alpha-tocopherol concentration in soybean seeds.

4. Development of high oleic acid gene-specific markers

   Two gene specific markers were developed for FAD2-1A and FAD2-1B responsible for high oleic acid production in soybean seeds.

   Further optimizations are ongoing, and markers will be available for screening for 2013 growing season.

5. Application of molecular markers for true F1 soybean identification

   A set of 15 SSR markers were used to detect polymorphisms among parental lines used for 2011 and 2012 crosses made in Inkerman experimental trials and at a greenhouse at the University of Ottawa. True F1 plants were verified with these markers and false F1 plants were discarded. These true F1 plants were advanced for higher generations.

6. Application of molecular markers for commercial variety sample discrimination

   A set of 10 highly polymorphic SSR markers across 10 soybean chromosomes were validated for soybean variety discrimination and variety identification tests. These 10 SSR markers were tested for identification of some commercial moldy soybean samples of DH530 exported to Japan. DNAs were extracted from seeds of both normal DH530 and some moldy samples. These markers were analyzed with both high concentration agarose gel (3%) and ABI 3730 DNA

Analyzer.

The moldy soybeans and DH530 have the same molecular patterns for all these 10 SSR markers. The results indicate that the commercial moldy soybeans exported to Japan are genetically identical to DH530 according to the DNA marker test. These moldy soybeans are indeed DH530, and are not contaminated with other moldy lines.

4. Next steps and references

1. Mutagenesis and gene targeted selection for cytokinin genes (Gm-IPT, and Gm-CKX genes) will be conducted to create desired mutants with robust root systems (drought tolerance) and better flowering and pod setting soybeans.

2. More targeted traits for marker development and selection

   Agronomic
   - Yield (cytokinin, brassinosteroids, circadian systems)
   - White mould resistance
   - Aphid resistance
   - Drought tolerance
   - Iron deficiency chlorosis
   - Maturity

   Food
   - Saponin A free
   - Low phytate
   - Low stachyose
   - Low raffinose
   - 7S and 11S proteins (7S α' null and 11S A4 null)
   - High oleic acid
   - Allergen free

References:

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M (2005). Science, 309:741-745

Bartrina I, E Otto, M Strnad, T Werner, T Schmulling (2011). Plant Cell, 23:69-80

Emery RJN, CA Atkins. (2006). In Handbook of Seed Science, (Amarjit S. Basra ed) Haworth's Food Products Press, Binghamton, New York. Pp. 63-93.

Emery RJN, L Leport, JE Barton, NC Turner, CA Atkins (1998). Plant Physiology, 117: 1515-1523.

Emery RJN, Ma Qifu, CA Atkins. (2000). Plant Physiology, 123: 1593-1605

Liu F.L., C. Jensen and M. Andersen, (2004). Annals of Botany, 94: 405–411

Morris, R.O. (1997). In Cellular and Molecular Biology of Plant and Seed Development. Ed BA Larkins IK Vasil. Kluwer Academic Publishers, pp. 117-149.

Nagel L, R Brewster, WE Riedell, N Reese. (2001). Annals of Botany, 88: 27-31

Quesnelle PE, Emery RJN (2007). Canadian Journal of Botany, 85: 91-103.

Powell A.F., A. Paleczny, H. Olechowski, and N. Emery, (2013). Plant Physiology and Biochemistry 64 :33 – 40

Zalewski W, P Galuszka, S Gasparis, W Orczyk, A Nadolska-Orczyk (2010). Journal of Experimental Botany, 61:1839-51

B (I). Funded Collaborators (Co-PI, AAFC, other federal scientists)

- Include the name of scientist / organization.

Zhiyong Zhang, Senior Scientist of Sevita International and Visiting Scientist at Trent University

B (II). Acknowledgement of non-funded collaborators (who provide support, e.g. access to other laboratory or other facilities and equipment input / advice / guidance / assistance, etc).

- For research supported by targeted funding programs (e.g. DIAP, Clusters, etc.) please list any collaborators who are receiving Contribution Vote 10 funds (e.g., university and industry collaborators). In addition, please list separately the participants who support your project but are not receiving any funding through the program.
- Include name of scientist / organization.

C. Variance Report (if applicable, describe how the work differs from the proposed research)

- Include changes to objectives and project work plan / budget, changes to the team, other constraints.

1. Gene numbers that were investigated dramatically increased from initially proposed three Gm-IPT genes and three Gm-CKX genes to actual 14 Gm-IPT genes, 17 Gm-CKX genes, and 22 Gm-LOG genes.

2. Markers for soybean improvement were done at Trent University instead of La Cite College.

D. Impact Assessment (if applicable, describe how the variance factors above will impact project continuation)

- Include changes to the objectives, changes to the project work plan / budget, changes to performance (i.e. meeting targets).

The workflow and accomplishments increased dramatically over that which was originally proposed.

The increased work that was acomplished, was in large part due to the efficient communication and collaborations between Trent and Sevita. In particular, Zhiyong Zhang was embedded at Trent as a Senior Scientist to assist Neil Emery to supervise the project and develop protocols for both Trent

graduate students and Sevita junior researchers. This model enabled the achievement of the original targets, and completed the additional gene cloning, sequencing, and marker development.

This critical work has laid the groundwork for important future studies to pinpoint and optimize the promotive effect of the cytokinins on soybean seed yield. While strongly exceeding the original goals, the characterization of so many CKX, IPT and LOG genes is a very valuable resource for soybean yield research. Directly, Trent has strong candidate markers (SNPs) in CK genes that are associated with seed expression and potentially yield. Moreover, there is now thorough knowledge of which CK genes are active in the seed, which may be subject to mutation to affect seed yield. The next steps would involve targeted mutations of these strategic genes.

E. Achievements (include only those related to this project)

- Include innovations, publications / conferences, technology transfer, capacity building, success stories, media, recognition and other outputs.

The discoveries for marker development (i.e. SNPs) and their validation have been transferred to Sevita International for incorporation into breeding selection protocols.

Two graduate students will present some of the results at national and international conferences. One meeting is the Canadian Society of Plant Biologists' Annual Meeting at Laval University from June 25 to June 28 2013, and the other meeting is Plant Biology 2013**,** Providence, Rhode Island July 20-24, 2013

F. Lessons learned (self-evaluation of project)

1. The soybean genome is very complex, and the whole soybean genome sequence annotation has been continually updated throughout the project. This project was proposed before the full soybean genome sequence was available, and the effective workload became much greater than originally anticipated, as the number of CK related genes mushroomed.

2. Soybean yield and many other agronomic traits are also complex; multiple factors must be taken into account for the data analysis and results interpretations.

Neil Emery		
PI Name	**Date**	**Signature**

Note: After completion and signature, this report must be provided to the appropriate Science Director for assessment. A PDF copy of this report will be sent to Science Operations by the Science Director's office along with the project assessment.

graduate students and Sevita junior researchers. This model enabled the achievement of the original targets, and completed the additional gene cloning, sequencing, and marker development.

This critical work has laid the groundwork for important future studies to pinpoint and optimize the promotive effect of the cytokinins on soybean seed yield. While strongly exceeding the original goals, the characterization of so many CKX, IPT and LOG genes is a very valuable resource for soybean yield research. Directly, Trent has strong candidate markers (SNPs) in CK genes that are associated with seed expression and potentially yield. Moreover, there is now thorough knowledge of which CK genes are active in the seed, which may be subject to mutation to affect seed yield. The next steps would involve targeted mutations of these strategic genes.

E. Achievements (include only those related to this project)

- Include innovations, publications / conferences, technology transfer, capacity building, success stories, media, recognition and other outputs.

The discoveries for marker development (i.e. SNPs) and their validation have been transferred to Sevita International for incorporation into breeding selection protocols.

Two graduate students will present some of the results at national and international conferences. One meeting is the Canadian Society of Plant Biologists' Annual Meeting at Laval University from June 25 to June 28 2013, and the other meeting is Plant Biology 2013, Providence, Rhode Island July 20-24, 2013

F. Lessons learned (self-evaluation of project)

1. The soybean genome is very complex, and the whole soybean genome sequence annotation has been continually updated throughout the project. This project was proposed before the full soybean genome sequence was available, and the effective workload became much greater than originally anticipated, as the number of CK related genes mushroomed.

2. Soybean yield and many other agronomic traits are also complex; multiple factors must be taken into account for the data analysis and results interpretations.

Neil Emery	May 29, 2013	
PI Name	**Date**	**Signature**

Note: After completion and signature, this report must be provided to the appropriate Science Director for assessment. A PDF copy of this report will be sent to Science Operations by the Science Director's office along with the project assessment.