

A genomic approach to nutritional, pharmacological and genetic issues of the faba bean (*Vicia faba*)

Prospects for genetic modifications

Heather Ray and Fawzy Georges*

Plant Biotechnology Institute; National Research Council of Canada; Saskatoon, SK Canada

Key words: *Vicia faba*, EST library, L-DOPA, defensin, metallothionein, vicine, divicine, convicine, isouramil, favism

Cultivated faba bean (*Vicia faba*) is widely used as human food, especially in Europe, Northern Africa and China. In view of its superior feeding value over field peas or other legumes, it is also widely used as animal feed for a variety of species. *V. faba* also contains medically important components such as 3,4-dihydroxyphenylalanine (levo-DOPA, L-DOPA), the principal treatment used for Parkinson disease patients. However, this species also contains several antinutritional components, including the pyrimidine glycosides vicine and convicine; phytates; and the sucrose galactosides including raffinose, stachyose and verbascose.

We have undertaken a genomic project to provide publicly available expressed sequence tag sequences (EST) prepared from early to mid developing embryo in an attempt to identify genes that are likely to be involved in the biosynthesis of L-DOPA and the vicine group of compounds. As initial examples of the utility of this approach, we describe the complete sequence of fabatin, new defensins, type 4 metallothioneins and a variety of other key genes which were identified in this EST library. No candidate sequences corresponding to the biosynthesis of L-DOPA or the vicine group could be identified at this early stage of seed development.

Do not distribute.

Introduction

The importance of legumes in agriculture, human consumption and animal nutrition is increasing exponentially due to the increasing world population and its need for proteins. Food legumes are considered the best substitute for meat in many parts of the world, where there is demand for alternate, non-animal protein sources. Legume crops have two distinctive traits: (1) their high protein content; and (2) their unique symbiotic ability to fix atmospheric nitrogen in the soil.

Faba bean (*Vicia faba*) (fava bean, broad bean, horse bean) is an important member of the legume family with highly useful characteristics. It is widely grown and consumed, especially in China, North African countries and parts of Europe and North and South America, and is served in a great variety of forms, mostly based on the immature or mature seed. For both humans and livestock, it provides high quality, lysine-rich proteins, carbohydrates and fibers. It is also rich in carotenoids, vitamins^{1,2} and essential minerals including iron, magnesium, potassium, zinc, copper and selenium. Faba beans have also been shown to have lipid-lowering effects and may also be a good source of antioxidants and chemopreventive factors.³

Particularly noteworthy in faba bean is the medically important component 3,4-dihydroxyphenylalanine (levo-dopa, L-DOPA).

The latter is the major ingredient in medicines used to treat Parkinson disease (PD) patients. The anti-PD effects of *V. faba* have been documented as superior to those of synthetic L-DOPA and considerably more lasting.^{4,5} Therefore, *V. faba* may be a crop well positioned for molecular pharming purposes, with the possibility of developing cultivars enriched in L-DOPA and other pharmaceuticals.

In common with numerous crop legumes, faba bean produces various antinutritional factors including raffinose series oligosaccharides, lectins and protease inhibitors, phytate, and tannins. Almost unique to faba bean are vicine and convicine, the causative agents of favism in many human populations. These antinutritional factors have limited its worldwide acceptance as a competitive food crop.

As an effective nitrogen fixing species, it is regarded as an excellent crop for soil amendment, which also provides high quality fodder and silage. In Canada, faba bean may have high potential for expansion, given its good tolerance of cool weather, and relative freedom from disease, insects and parasites compared to many areas where it is more commonly grown.

Elsewhere around the world, faba bean is grown as a seed crop mostly along the west coast of the United States. In South America faba bean is also of importance as a food crop especially in the Andean region. Most significantly, the crop is very

*Correspondence to: Fawzy Georges; Email: fawzy.georges@nrc-cnrc.gc.ca

Submitted: 02/07/10; Revised: 03/26/10; Accepted: 03/26/10

Previously published online: www.landesbioscience.com/journals/gmcrops/article/11891

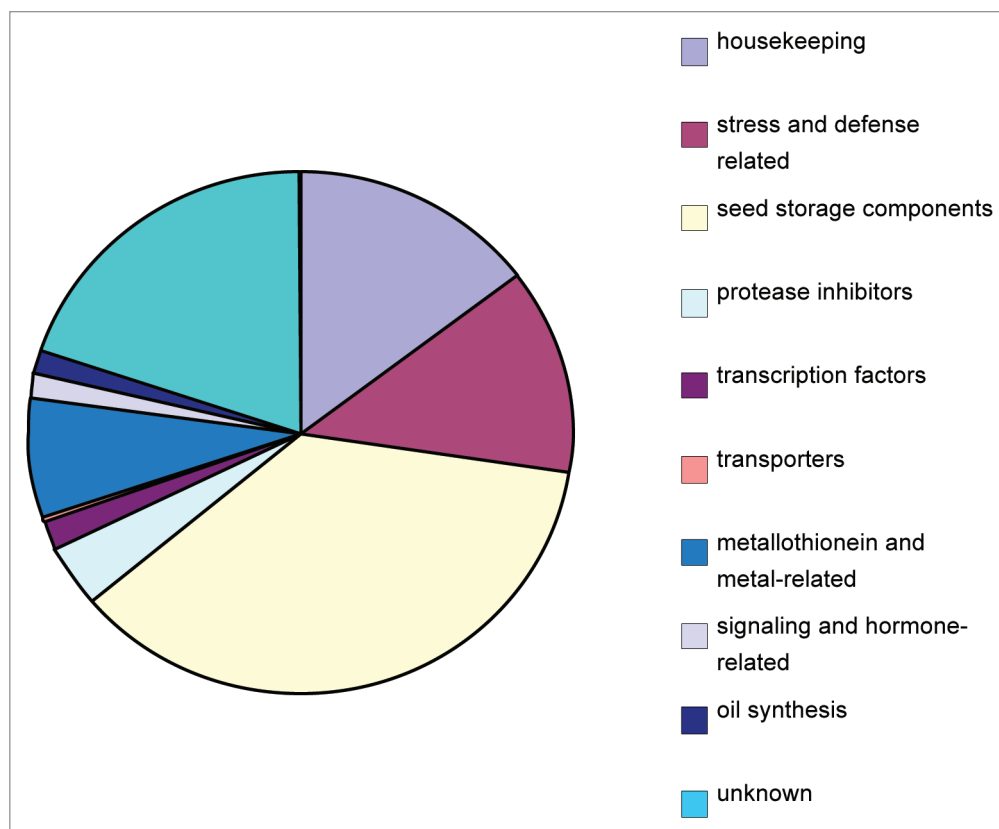


Figure 1. Summarized description of gene function for faba bean library.

DO NOT distribute.

widely distributed in the Mediterranean region, the Nile valley, Ethiopia, Central Asia and Northern Europe.

In proportion to its potential, this species has received very limited attention in the form of plant breeding programs or molecular studies, such as genomics and gene expression analyses. It has not been possible to cross faba bean with other species and its progenitor species, if extant, is unknown.^{6,7} It is also difficult to transform, although methods have been published.⁸⁻¹⁰ There is an almost complete absence of published sequence data for this species, despite the advantages of such data for breeding efforts and identifying particular genes of high interest. Prior to this project, fewer than 400 gene sequences of *V. faba* have been submitted to GenBank.

EST libraries offer the advantage of rapid identification and cataloguing of key tissue-specific or stage-specific genes, particularly for a species such as *V. faba*, which has very large chromosomes and genome size, and has been minimally mapped. The establishment of a broad genomic database for *V. faba* is essential both to enable marker-assisted selection for breeding and to allow full exploitation of its potential benefits in health and agriculture. Knowledge of the dynamics of temporal and spatial gene expression in this valuable member of the legume family is important for effective trait- and seed-quality improvements. Varshney et al.¹¹ have recently identified faba bean as among the “orphan legume” crops much in need of a genomics research effort; of this group, it is almost the only one to have received no such attention.

In the initial phase of a faba genomics program, we describe the development of an EST library of about 5,000 sequences from early to mid developing embryo, and present some of the results of preliminary analyses. We also examine some of the various benefits and limitations associated with this crop, confirming the need for a comprehensive genomics study of *V. faba*.

Results and Discussion

One of our initial objectives was to expand the knowledge of seed proteins of *V. faba*; a second objective was to screen for stage-specific transcripts likely to be related to vicine, convicine and L-DOPA biosynthesis. Therefore, the choice of tissue in this study was based on the earlier observation that the accumulation of vicine and convicine in developing faba beans was highest in developing cotyledons (moisture content about 80%).¹²

The distribution of metabolic classes among library genes appears in **Figure 1**, while **Table 1** displays some of the most frequent groups of transcripts. The legumins, albumins and glycinins predominate among storage proteins, while a wide range of late-embryo-abundant and seed maturation proteins occurs. Convicilin in significant amounts was detected, but vicilin expression was much lower. Unexpectedly, given that the oil content of mature faba seed is only 1–3%, large amounts of oleosin, caleosin and lipid transfer protein transcripts were also present. Also present in large numbers are fabatin, defensin and protease inhibitor transcripts, together with large numbers of metallothionein-like

transcripts, discussed further below. For each class, the approximate number of types probably represents a minimal estimate, as contigs may contain up to 5% non-identical nucleotides. Numerous full-length or near-full-length genes, some of which appear in Table 2, were present in the library. They include the albumins, desiccation and dormancy related genes, peroxiredoxin and superoxide dismutase genes, and individual regulatory and kinase genes which are present in relatively high abundance.

In addition, a very wide range of the transcripts expected in an actively growing tissue are found, as shown in Figure 1. The detailed function for the majority of these genes are unknown, as the phylogenetic distance to *Pisum sativum*, the nearest congener with a reasonable number of GenBank entries, is enough that numerous subtle differences would be expected.

L-DOPA. L-DOPA (Fig. 2) is found in several tissues of faba bean¹⁴ and may be highest in young pod material (6–7%)¹² and seedlings (2% in whole 9-day seedling, chiefly in the axis).¹⁵ It is present in moderate amounts in young developing seed but decreases gradually as seed matures,¹⁶ suggesting that it is synthesized or possibly imported only at an early stage. This species is almost unique among plants in accumulating L-DOPA.

The utility of L-DOPA to the plant is unclear, although allelopathic effects¹⁷ and toxicity to insects¹⁸ have been suggested. Further exploration of this area might be valuable, particularly as L-DOPA is a key neurochemical.

In humans, L-DOPA can cross the blood-brain barrier into the brain where it is converted to dopamine. Pharmaceutically produced L-DOPA controls the symptoms of PD but gives somewhat inconsistent control and may produce hallucinations. However, eating young sprouts, seeds or pods of faba bean is reported to relieve Parkinson's symptoms without the inconsistent control of the pharmaceutical form.^{4,5} While this rests on rather few controlled studies, and may be due solely to the slower release of active compounds through digestion, it is possible that additional anti-Parkinsonian factors may await discovery in faba bean.

The synthetic pathway of these compounds is not fully understood. In faba bean, tyrosine is converted to L-DOPA by a hitherto unidentified oxidase. As L-DOPA and tyrosine differ only by one hydroxyl group on the phenol ring, the chemistry looks straightforward. However, an enzyme comparable to tyrosine hydroxylase of mammals is not known from plants.

Tyrosinase (polyphenol oxidase) is known to add a second hydroxyl group to monohydroxylated phenol rings, including tyrosine, but this is normally followed by their immediate conversion by the same enzyme to quinone groups. The two activities have been considered inseparable in plant polyphenol oxidases, but a fungal polyphenol oxidase has been shown to have high tyrosine hydroxylase activity with very low quinone-forming activity,¹⁹ while in plant species that produce betalains, a tyrosine hydroxylation followed by cyclization (rather than quinone formation) is the first step of betalain synthesis, and in *Portulaca*, tyrosine hydroxylases have been partly characterized.^{20,21} In humans, a tyrosinase may also contribute to L-DOPA synthesis.²² It is possible that a faba bean form of this enzyme may have similar characteristics.

Table 1. High-frequency transcripts in embryo tissue of *V. faba*

Class	Approximate number of types	% of total ESTs
Albumins	10	7.2
Seed maturation proteins	18	7.3
Oleosins	5	2.7
Caleosin	1	0.6
Lipid transfer proteins 4	5	1.2
Embryo proteins	2	2.9
Late embryo abundant proteins	8	1.5
Legumin As	7	3.1
Fabatins	2	1.7
Glycinin subunits 7	2	1.3
Convicilins	4	0.8
Dehydrins 2	4	1.7
Trypsin/chymotrypsin inhibitors	4	3.8
Metallothionein-like proteins	4	6.6
Defensins and defensin-like proteins	9	8.5

An alternative may be that L-DOPA is immediately removed from the enzyme following the first reaction and sequestered from it by an unknown mechanism. At least some L-DOPA resides in the leaf vacuole, kept in a reduced state by ascorbic acid.²³ Two amide-linked conjugates of jasmonic acid, *N*-[(3*R*,7*R*)-(-)-jasmonyl]-(*S*)-dopa and *N*-[(3*R*,7*R*)-(-)-jasmonyl]-dopamine have been isolated from flowers of *V. faba*.²⁴ This, together with the finding that crude enzyme preparations from the same source were shown to effectively hydroxylate the known *N*-(jasmonyl) tyrosine conjugate, suggests jasmonic acid conjugates as possible intermediates in the biosynthesis of L-DOPA, or catabolic derivatives of jasmonic acid, in *V. faba*. However, no candidate transcripts related to the formation of this type of conjugates or their hydrolysis could be found at this stage of embryo development.

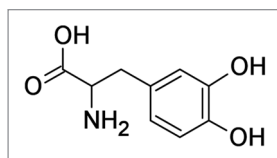
Three polyphenol oxidase sequences of *V. faba* have previously been isolated.²⁵ In leaf and pod libraries under development, we have identified additional sequences (results not shown). However, contrary to expectations, none were found in this early-stage embryo library. This may point either to the involvement of a different type of enzyme activity, or to the import of pre-synthesized L-DOPA into the embryo.

Vicine and convicine. Vicine and convicine are simple glycosides derived from the pyrimidine biosynthetic pathway, and occur in substantial amounts in some tissues of faba bean, particularly the immature seed, and drop considerably in the mature seed, rising again in the young seedling. In immature seed, the concentration of vicine averaged 1.3% in a set of nine lines, while that of convicine averaged 0.4%.²⁶ The biological effects of vicine and convicine are believed to be mediated through their aglycones, divicine and isouramil, respectively.²⁷⁻²⁹ For example, divicine has been shown to dramatically reduce erythrocyte survival in vitro.³⁰ When consumed, vicine and convicine are converted to the active hemolytic aglycones by β -glucosidase (Fig. 3). Thus, vicine and convicine constitute a rare class of antinutritional factors which is

Table 2. Some full length and near-full-length genes moderately to highly expressed in early to mid embryo developmental stages

Gene	Identity	Frequency, %	Length
REHY MEDTR	1-Cys peroxidoredoxin	0.29	full
Q8GUR9 PEA	Thioredoxin h	0.06	full
SODCP PEA	Superoxide dismutase [Cu-Zn]	0.10	full
Q5DWE8 PEA	Superoxide dismutase [Cu-Zn]	0.08	full
Q76MF3 TOBAC	Calmodulin NtCaM3	0.04	full
O04117 PEA	Dehydrin 2	1.0	full
ALB1A PEA	Albumin-1 A	0.08	full
ALB1E PEA	Albumin-1 E	73	full
Q8S4Q2 9FABA	Ethylene-responsive transcriptional coactivator-like protein; multiprotein bridging factor 1c	1.52	full
Q9LQP1 ARATH	Putative receptor serine/threonine kinase; glyoxalase I family	0.08	full
A9PIK1 POPTR	Putative uncharacterized protein; multiprotein bridging factor 1b	0.06	full
Q8LKG1 PEA	Dormin3	0.90	full
O22611 PEA	Dormancy-associated protein	0.25	full
Q2XSK1 GLYSO	Seed maturation protein	0.10	full
AT5G11340	GCN5-related N-acetyltransferase (GNAT) family	0.06	full
AT5G01300	phosphatidylethanolamine-binding family	0.19	full
Q2Q4X9 MEDTR	seed maturation protein	0.71	near
AT1G51200	zinc finger (AN1-like) family	0.08	full
AT2G32090	lactoylglutathione lyase family	0.04	full
AT2G46540	Similar to fiber protein FB11	0.04	near
Q6J338	Copper chaperone	0.04	near
Q9S7N8 SOYBN	Seed maturation protein PM21	0.23	near
AT4G10270	wound-responsive family	0.08	full

The listed genes are all relatively short. Genes described in the text, housekeeping and those related to uncharacterised genes are excluded.

**Figure 2.** L-DOPA (3,4-Dihydroxyphenylalanine).

nearly exclusive to the genus *Vicia*. They are the causative agents of a medical syndrome known as favism, an inherited disorder to which individuals who are deficient in glucose-6-phosphate dehydrogenase (G6PD) are most susceptible.³¹ The enzyme G6PD functions to maintain adequate supplies of reduced glutathione, which in turn reduces these oxidants and renders them inactive.^{32,33} G6PD deficiency, a recessive sex-linked trait, is one of the most common human enzyme deficiencies worldwide and about 400 million people are affected by this enzymopathy on a global scale.

Although vicine and convicine have not been considered general antinutritional factors for livestock, chickens fed on high-vicine vs. low-vicine feed exhibit a lower rate of gain³⁴ and reduced egg production and quality.³⁵

It is assumed that uracil is the progenitor compound of vicine and convicine, but the pathway has not been characterized to

date. The unknown provenance of these factors has not prevented plant breeders from identifying low-vicine and convicine lines carrying the mutation *Vc*, which reduces the amount of both compounds by 90–95%.^{1,36} *Vc* must be a step in the common pathway and may be a regulatory gene, but the nature of the mutation is not known. This mutation is present in many recently released varieties. The relative proportions of vicine and convicine vary significantly among varieties,^{1,26} but varieties with only one of the two compounds have not been identified, which suggests synthetic interconversion between the two forms.

While vicine and convicine may be presumed to have some function in faba bean, the *Vc* mutant lacks obvious deleterious effects in terms of protein content, seed yield, and disease susceptibility.³⁷ However, antifungal effects³⁸ and deleterious effects on generalist insects³⁹ have been noted. In addition, vicine and convicine appear to have some antimalarial effect, mediated by oxidative damage.⁴⁰

Whether through conventional breeding, mutagenesis or genetic engineering, identification of genes involved in the biosynthesis of these antinutritional factors is essential for developing efficient programs for their removal. No obvious candidates for such genes could be detected in the present EST library, suggesting that the accumulation of vicine and convicine in cotyledons at this developmental stage may be due to translocation

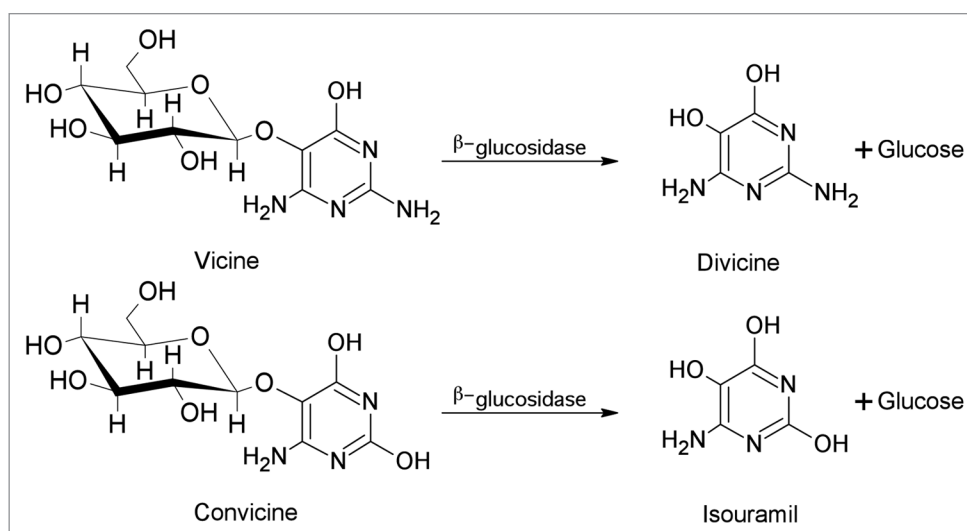


Figure 3. Vicine and convicine are the β -glucosyl derivatives of divicine and isouramil respectively, which are the actual causative agents of favism.

rather than de novo synthesis. They are likely to be synthesized at least partly in maternal tissue, in view of the observation that maternal effect occurs in crosses.³⁶ EST libraries from maternal tissues would, therefore, be advantageous. Elucidation of this synthetic pathway might contribute, in addition, to characterizing the unknown synthetic pathways of several other unusual pyrimidines found in various legume species.⁴¹

Trypsin inhibitors and lectins. Faba beans contain trypsin inhibitors and lectins which reduce seed digestibility, in common with many bean and pulse crops. However, the average trypsin/chymotrypsin inhibitor level of several faba varieties is about 10% of levels in soybean.⁴² Similarly, faba bean lectin proteins are also found at low levels and, therefore, neither issue is considered of critical importance in this species. Transcripts corresponding to trypsin and chymotrypsin inhibitors appear to be present in higher abundance than the lectin transcripts at the 15 DAF developmental stage. Based on the EST data, transcripts corresponding to trypsin and chymotrypsin inhibitors occurred at 3.7% compared to 0.3% for lectin-related transcripts. Most strongly expressed is a Bowman-Birk type inhibitor which is highly similar to the pea (*Pisum sativum*) trypsin/chymotrypsin inhibitor (87% AA identity to GenBank Accession CAB92980). A Kunitz type inhibitor was also observed at much lower expression levels; it is most similar to a chickpea (*Cicer arietinum*) gene (76% AA identity to GenBank Accession CAH61462). These expression levels appear high, considering their relatively minor importance in this species.

Oligosaccharides of the raffinose family. Legumes are generally rich in members of the raffinose family of oligosaccharides (RFOs) which includes, in addition to raffinose, stachyose, verbascose and higher oligomers. RFOs are involved in the storage and transport of carbon and are believed to play a protective role of cell membranes as the seed desiccates and matures, and possibly as it germinates.⁴³ In some species, they appear to be implicated in abiotic stress tolerance acquisition.⁴⁴⁻⁴⁶

RFOs are synthesized through sequential addition of galactosyl residues, which are donated through transglycosylation from galactinol (*myo*-inositol α -D-galactoside), to sucrose. The resulting sucrose galactosides are considered antinutritional agents for humans and monogastric animals. They are not readily digested, due to the absence of α -galactosidase activity in the alimentary canal, leading to anaerobic digestion and the generation of methane gas. They occur to a total of 1.4–6.2% of dry weight in faba bean seed, verbascose being the largest component¹ (Georges et al., unpublished results). These levels are somewhat higher than those identified in common bean, *Phaseolus vulgaris*.⁴⁷

The biosynthesis of RFOs depends largely on the availability of the key intermediate, galactinol, which is assembled by the action of galactinol synthase (GolS), using UDP-galactose and *myo*-inositol. In the current EST library, transcripts corresponding to *GolS* with high homology to the Arabidopsis *GolS1* isoform (At2g47180) were detected. The corresponding translated sequence was also homologous to a putative GolS protein from pea (Q9XGG4). The level of this transcript appears to be low (0.12%), suggesting that accumulation of these sugars is minimal at this stage.

Limiting the availability of the GolS enzyme or one of its substrates, *myo*-inositol, in faba bean might lead to reduced levels of the RFO sugars and improvement of the nutritional quality of the bean. A similar approach has been reported in *Brassica napus*.¹⁵ We have used EST data to isolate full-length *V. faba GolS* (Georges et al., unpublished results).

Phytates. The chief storage form of phosphorus in mature faba bean seed is phytic acid (phytate). The average phytate level of a set of faba bean lines was found to be 1.6%.⁴² In the germinating seed this is degraded by phytase, providing phosphorus for the seedling metabolic needs. The amounts of stored phytate in cultivated plants may be surplus to requirements.

Phytate phosphorus is largely unavailable to monogastric livestock because they lack sufficient phosphatase activity to

MEKKAVAALSFLFLVLFVAQEIVVTEAR	TCEHLADTYRGPCLTDASCDDHCNKKAHLISGTCHN-YKCFCTQNC	def1	5.6%
MEKKAVAALSFLFLVLFVAQEIAVSEAK	TENLSDTFKGPCIPDGNCKNHCKNNEHLLSGRCRDDFRWCTRNC	def2	2.5%
MEKKSVAALSFLFLVLFVAQEIVVTEAR	TCEHLADTYRGPCLTNASCDDHCNKKAHLISGTCHN-YKCFCTQNC	def3	0.6%

Figure 4. Amino acid alignment of *V. faba* defensins. The end of the signal peptide is marked by a gap, and the cysteines are underlined.

---LLGRCKVKSNRNFNGPCLTDTHCSTVCRGEGYKGGDCHGFRRRCMCLC
 MERKTLSTFTFMLFLLLVDVSVKTSEALLGRCKVKSNRNFNGPCLTDTHCSTVCRGEGYKGGDCHGFRRRCMCLC

Figure 5. Amino acid sequence of fabatin (47 aa) and amino acid sequence derived from contig containing gene for this protein (74 aa). The cysteines are underlined.

hydrolyze the phosphate groups from the phytate core structure (*myo*-inositol). Thus, the inability of the animals to degrade phytate results in the secretion of large amounts of total phosphate, which leads to environmental pollution, especially in areas of intensive pig production. Moreover, the ability of phytate to chelate positively charged species including proteins and essential minerals such as iron, magnesium and zinc, prevents their assimilation and, in a grain and legume based diet, can result in symptoms of malnutrition and reduced productivity.⁴⁸ Therefore, for improved food and feed quality, phytate reduction by alteration of its biosynthetic pathway is advantageous. This may be achieved through conventional breeding programs, mutagenesis or by transgenic approaches.

myo-Inositol is a shared substrate, common to both phytate and RFO biosynthetic pathways. Decreasing the supply of *myo*-inositol in seeds can affect the accumulation of both antinutritional factors, as found in similar experiments with *B. napus* (Georges et al., unpublished results). The supply of *myo*-inositol in developing seeds can either be through de novo synthesis or catabolic salvage from other *myo*-inositol-containing substrates. The de novo synthesis of *myo*-inositol is carried out by the enzyme *myo*-inositol phosphate synthase (MIPS). In the present EST library, the presence of *MIPS* transcripts (0.08%) indicates that de novo synthesis of *myo*-inositol is occurring during this stage of seed development, although at a low level. We have used EST data to isolate full-length *V. faba MIPS* (Georges et al. unpublished results).

Metallothioneins. The metallothioneins constitute a class of proteinaceous metal chelators. The metallothioneins (Mt) or metallothionein-like genes of faba bean are highly expressed at this early stage of embryo development. In the EST library, Mt-like transcripts comprised 4.2% of the total, some of which have been characterized previously, while others are new (Fig. 6A). One of these genes is identical to X77254, highly expressed in leaf trichomes;⁴⁹ another is identical to a type 1 Mt, X91077, expressed in several tissues.⁵⁰ Very highly expressed (4%) is a protein with the characteristic cysteine distribution of a type 4 Mt (Accession EU920049; Fig. 6B), a class with seven C-X-C dyads.⁵¹ While some plant species accumulate fairly high levels of type 4 Mt in seeds in unstressed conditions,⁵¹ its role is unclear. An Arabidopsis type 4 Mt, transformed into yeast, appeared to confer tolerance of copper to the same degree as other forms of Mt, but to confer greater tolerance of zinc.⁵² However, the wheat *Ec*

gene, a type 4 Mt, is reported to have low affinity for zinc ions, and is postulated as a zinc trafficking protein, particularly during germination.⁵³

In mammalian systems, in addition to their role as antioxidants, Mt(s) also appear to play a regulatory role. For example, by binding and releasing zinc, they may regulate zinc levels within the body and play a role in zinc signaling within cells.^{54,55} The ability of Mt(s) to interchange their zinc with zinc-finger proteins also suggests a contributory role of Mt(s) to zinc-dependent processes involved in gene expression.⁵⁶

Although *V. faba* does not appear to be generally among the most metal tolerant plant species, it is reported to hyperaccumulate lead^{57,58} and has been suggested as a suitable crop for phytoremediation.⁵⁹

Disease resistance and defense. Disease is a major issue for growers of faba bean,⁶⁰ and various programs attempt to associate markers of different kinds with known or candidate disease resistance genes.⁶ Of defense-related compounds, forms of defensin and fabatin formed a substantial part of the total population of transcripts in the present EST library (8.5 and 1.7%, respectively). Three major forms of embryo-expressed defensin, which we term *Def1*, 2 and 3, were identified (Fig. 4). *Def1* and *Def2* were both encoded by more than one nucleotide sequence (*Def1* by GenBank Accession EU920044 and EU920047, *Def2* by EU920045 and EU920046), while *Def3* was encoded by EU920048. Additional minor forms were detected. These are typical defensins, with a highly conserved signal peptide sequence and eight conserved cysteine residues. The amino acid sequences are approximately 90% identical to an alfalfa defensin (GenBank Accession AAT66096). However, their antimicrobial activity has not been established.

Fabatins are also highly expressed in the embryo tissue (1.8% of the EST population). Fabatins are proteins with antimicrobial properties and, like the defensins, have a high proportion of cysteine in a slightly different arrangement which is identical to the γ -thionins of wheat and barley.⁶¹ We identified a fabatin (GenBank Accession EU920043) which is almost identical to fabatin 1 and 2 (GenBank Accessions P81456 and P81457),⁶¹ and found that its sequence included a signal peptide-like sequence in a longer open reading frame, bringing it to AA74 residues (Fig. 5). It is, therefore, established as a defensin or protease inhibitor, which is similar to a probable protease inhibitor of potato (GenBank Accession P20346),⁶² while its putative signal peptide

A

MSCCGGNCGCGSSCKCGSGCGGCKMYADLSYTESTTSETLIMGVGSEKAQYESAEMGAENDGCKCGANCTCNPCTCK

B

MADTGRVVVCDNRCGCTVPCAGGSTCRCTSSEGGARTDHTTPCGEHCECNPNCSRTVAAGSGCRCDASCTCASCRT

Figure 6. Metallothionein sequences predominant in *V. faba*. (A) Established Mt, representing about 0.4% of transcripts. (B) New Mt, representing about 4% of transcripts. Note the five cysteine dyads in (A) and the seven in (B), underlined.

sequence most resembles that of a putative defensin of *Vigna radiata* (GenBank Accession AAR08912).

Recently, small cysteine-rich polypeptides such as the defensins and protease inhibitors have been found to be much more abundant in plant genomes than was evident in earlier genomic efforts.⁶³ The relative abundance of corresponding transcripts in the present EST library is in further accord. However, their role is not well understood.

Faba bean as a GM crop. Faba bean is a crop poised for unique genetic modification opportunities which are not only extendable to trait improvement but also issues related to human health and molecular pharming. This species is suited for pharming in some regards, particularly for its native pharmaceutical compounds. However, consistent with the very modest level of molecular research it has received, reports on faba transformation have been scarce. Tissue culture and regeneration protocols have recently been described for this species,⁶⁴ but very few studies of genetically transformed *Vicia faba* have been reported, and productivity to date is low.⁸⁻¹⁰

In order to improve the nutritional and nutraceutical values of faba bean and enhance its resistance to biotic and abiotic stresses, traditional breeding programs could advantageously be supplemented by functional genomics and gene transfer studies. The development of larger-scale genomic databases will provide excellent prospects for advancing this species quickly from a status of almost no genomic information to a state where substantial information on numerous aspects of its biology is available, so that faba will be able to fulfill its potential as an important nutraceutical legume which is also highly suitable as food, feed and nitrogen-providing crop.

Materials and Methods

Plant material and library preparation. Faba bean cv. Exhibition Long, a selection from Windsor Broad, was grown in a growth

cabinet and developing embryos collected at the 80% moisture level (15 days after flowering, DAF),¹² frozen in liquid nitrogen, ground, and RNA prepared using an RNeasy Plant RNA kit (Qiagen). RNA was then treated with DNase (Invitrogen) according to manufacturer's instructions, and cDNA prepared. A near-full-length library was prepared using Creator Smart cDNA library construction kit (Clontech), tested, and expanded to about 5,000 clones. The average length of inserts was approximately 700 bp. Sequence data has been submitted to GenBank (GI:219212932 to GI:219282595). Data was analyzed using standard analytic programs and assembled into contigs with a minimum of 95% nucleotide identity.

Seed components. Phytate and sugar components were analyzed using established HPLC and GC methods.¹³

Future Developments and Conclusions

In order to further understand the molecular dynamics that make faba bean unique among other legume crops, additional EST libraries, from a selected set of tissues and developmental stages, are now being prepared and sequenced. Since prior knowledge of the genes of *V. faba* is scant, such an assembly of EST libraries is expected to produce sequences of significant interest which will prove useful both for breeding purposes and molecular research. This might involve, for example, increasing L-DOPA or eliminating vicine and convicine for human consumption.

Acknowledgements

We are grateful for the technical assistance of Yurdagul Ferhatoglu, Albana Zeko and Cheryl Bock, the PBI DNA Sequencing Unit and the Bioinformatics Unit. Funding was provided in part by the Natural Products Genomics consortium of NRC. This is NRC publication 50114.

References

- Duc G, Marget P, Esnault R, LeGuen J, Bastianelli D. Genetic variability for feeding value of faba bean seeds (*Vicia faba*): Comparative chemical composition of isogenics involving zero-tannin and zero-vicine genes. *J Agric Sci* 1999; 133:185-96.
- Randhir R, Shetty K. Microwave-induced stimulation of L-DOPA, phenolics and antioxidant activity in faba bean (*Vicia faba*) for Parkinson's diet. *Process Biochem* 2004; 39:1775-84.
- Madar Z, Stark AH. New legume sources as therapeutic agents. *Br J Nutr* 2002; 88:287-92.
- Rabey JM, Vered Y, Shabtai H, Graff E, Korczyn AD. Improvement of Parkinsonian features correlate with high plasma levodopa values after broad bean (*Vicia faba*) consumption. *J Neurol Neurosurg Psychiatry* 1992; 55:725-7.
- Apaydin H, Ertan S, Ozekmekci S. Broad bean (*Vicia faba*)—a natural source of L-Dopa—prolongs "on" periods in patients with Parkinson's disease who have "on-off" fluctuations. *Move Disord* 2000; 15:164-6.
- Torres AM, Román B, Avila CM, Satovic Z, Rubiales D, Sillero JC, et al. Faba bean breeding for resistance against biotic stresses: Towards application of marker technology. *Euphytica* 2006; 147:67-80.
- Duc G, Bao S, Baum M, Redden B, Sadiki M, Suso MJ, et al. Diversity maintenance and use of *Vicia faba* L. genetic resources. *Field Crops Res* 2008; 115:270-8.
- Böttinger P, Steinmetz A, Schieder O, Pickardt T. Agrobacterium-mediated transformation of *Vicia faba*. *Mol Breed* 2001; 8:243-54.
- Hanafy M, Pickardt T, Kiesecker H, Jacobsen H-J. Agrobacterium-mediated transformation of faba bean (*Vicia faba* L.) using embryo axes. *Euphytica* 2005; 142:227-36.

10. Yan H, Yan H, Li G, Gong W, Jiao H, Chen H, Ji M. Expression of human cytomegalovirus pp150 gene in transgenic *Vicia faba* L. and immunogenicity of pp150 protein in mice. *Biologicals* 2010; 38:265-72.
11. Varshney RK, Close TJ, Singh NK, Hoisington DA, Cook DR. Orphan legume crops enter the genomics era! *Current Opin In Plant Biol* 2009; 12:202-10.
12. Burbano C, Cuadrado C, Muzquiz M, Cubero JL. Variation of fabaism-inducing factors (vicine, convicine and L-DOPA) during pod development in *Vicia faba* L. *Plant Foods Human Nutr* 1995; 47:265-74.
13. Bock C, Ray H, Georges F. Downregulation of galactinol synthase in oilseed *Brassica napus* leads to significant reduction of antinutritional oligosaccharides. *Botany* 2009; 87:597-603.
14. Longo R, Castellani A, Sberze P, Tibolla M. Distribution of L-Dopa and related amino acids in *Vicia*. *Phytochem* 1974; 13:167-71.
15. Goyoaga C, Burbano C, Cuadrado C, Varela A, Guillaumon E, Pedrosa MM, Muzquiz M. Content and distribution of vicine, convicine and L-DOPA during germination and seedling growth of two *Vicia faba* L. varieties. *Eur Food Res Technol* 2008; 227:1537-42.
16. Hill-Cottingham DG, Purves JV. Changes during development in the free amino acid constituents of fababeans (*Vicia faba* L.) plants. *Plant Soil* 2006; 75:435-42.
17. Nishihara E, Parvez MM, Araya H, Kawashima S, Fujii Y. L-3-(3,4-Dihydroxyphenyl)alanine (L-DOPA), an allelochemical exuded from velvetbean (*Mucuna pruriens*) roots. *Plant Growth Regul* 2005; 45:113-20.
18. Rehr SS, Janzen DH, Feeny PP. L-Dopa in Legume Seeds: A Chemical Barrier to Insect Attack. *Science* 1973; 181:81-2.
19. Hernández-Romero D, Sanchez-Amat A, Solano F. A tyrosinase with an abnormally high tyrosine hydroxylase/dopa oxidase ratio. *FEBS J* 2006; 273:257-70.
20. Steiner U, Schliemann W, Böhm H, Strack D. Tyrosinase involved in betain biosynthesis of higher plants. *Planta* 1999; 208:114-24.
21. Yamamoto K, Kobayashi N, Yoshitama K, Teremoto S, Komamine A. Isolation and purification of tyrosine hydroxylase from callus cultures of *Portulaca grandiflora*. *Plant Cell Physiol* 2001; 42:969-75.
22. Mastore M, Kohler L, Nappi AJ. Production and utilization of hydrogen peroxide associated with melanogenesis and tyrosinase-mediated oxidations of DOPA and dopamine. *FEBS J* 2005; 272:2407-15.
23. Takahama U. Hydrogen peroxide scavenging systems in vacuoles of mesophyll cells of *Vicia faba*. *Phytochem* 1992; 31:1127-33.
24. Kramell R, Schmidt J, Herrmann G, Schliemann W. *N*-(Jasmonoyl)tyrosine-Derived Compounds from Flowers of Broad Beans (*Vicia faba*). *J Nat Prod* 2005; 68:1345-9.
25. Cary JW, Lax AR, Flurkey WH. Cloning and characterization of cDNAs coding for *Vicia faba* polyphenol oxidase. *Plant Mol Biol* 1992; 20:245-53.
26. Hussein L, Motawel H, Nassib A, Khalil S, Marquardt R. The complete elimination of vicine and convicine from the faba beans by combinations of genetic selection and processing techniques. *Qual Plant Plant Foods Hum Nutr* 1986; 36:231-42.
27. Mager J, Glaser G, Razin A, Izak G, Bien S, Noam M. Metabolic effects of pyrimidines derived from faba bean glycosides on human erythrocytes deficient in glucose-6-phosphate dehydrogenase. *Biochem Biophys Res Comm* 1965; 20:235-40.
28. Arbid MSS, Marquardt RR. Hydrolysis of the toxic constituents (Vicine and convicine) in fababeans (*Vicia faba* L.) food preparations following treatment with β -glucosidase. *J Sci Food Agric* 1985; 36:839-46.
29. Winterbourn CC, Cowden WB, Sutton HC. Auto-oxidation of dialuric acid, divicine and isouramil. Superoxide dependent and independent. *Biochem Pharmacol* 1989; 38:611-8.
30. McMillan DC, Schey KL, Meier GP, Jollow DJ. Chemical analysis and hemolytic activity of the faba bean aglycon divicine. *Chem Res Toxicol* 1993; 6:439-44.
31. Duke JA. *Handbook of legumes of world economic importance*. Plenum Press, New York. 1981; 199-265.
32. Jacobasch G, Rapoport SM. Hemolytic anemias due to erythrocyte enzyme deficiencies. *Mol Aspects Med* 1996; 17:143-70.
33. McMillan DC, Bolchoz LJ, Jollow DJ. Favism: effect of divicine on rat erythrocyte sulfhydryl status, hexose monophosphate shunt activity, morphology and membrane skeletal proteins. *Toxicol Sci* 2001; 62:353-9.
34. Arese P, Duc G, Lessire M, Margot P. Low vicine-convicine and zero tannin "Fevita" faba beans. *Grain Legumes* 2007; 48:16-7.
35. Muduuli DS, Marquardt RR, Guenter W. Effect of dietary vicine and vitamin E supplementation on the productive performance of growing and laying chickens. *Br J Nutr* 1982; 47:53-60.
36. Duc G, Sixdenier G, Lila M, Furtoss V. Search of genetic variability for vicine and convicine content in *Vicia faba* L. A first report of a gene which codes for nearly zero-vicine and zero-convicine contents. In: Huisman J, Van der Poel AFB, Liener IE, eds. Recent advances of research in antinutritional factors in legume seeds. Wageningen: Academic Publishers 1989; 305-13.
37. Duc G. Faba bean (*Vicia faba* L.). *Field Crops Res* 1997; 53:99-109.
38. Bjerg B, Heide M, Knudsen JCN, Soerensen H. Inhibitory effects of convicine, vicine and dopa from *Vicia faba* on the in vitro growth rates of fungal pathogens. *Zeitschrift fuer Pflanzenkrankheiten und Pflanzenschutz* 1984; 91:483-7.
39. Desroches P, El Shazly E, Mandon N, Duc G, Huignard J. Development of *Callosobruchus chinensis* (L.) and *C. maculatus* (E) (Coleoptera: Bruchidae) in seeds of *Vicia faba* L. differing in their tannin, vicine and convicine contents. *J Stored Prod Res* 1995; 31:83-9.
40. Clark IA, Chaudhri G, Cowden WB. Some roles of free radicals in malaria. *Free Radical Biol and Med* 1989; 6:315-21.
41. Kafer C, Zhou L, Santoso D, Guirgis A, Weers B, Park S, Thornburg R. Regulation of pyrimidine metabolism in plants. *Front Biosci* 2004; 9:1611-25.
42. Makkar HPS, Becker K, Abel H, Pawelzik E. Nutrient contents, rumen protein degradability and antinutritional factors in some colour- and white-flowering cultivars of *Vicia faba* beans. *J Sci Food Agric* 1997; 75:511-20.
43. Obendorf RL. Oligosaccharides and galactosyl cyclitols in seed desiccation tolerance. *Seed Sci Res* 1997; 7:63-74.
44. Tajiri T, Ohsumi C, Iuchi S, Seki M, Kasuga M, Kobayashi M, et al. Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. *Plant J* 2002; 29:417-26.
45. Pennycooke JC, Jones ML, Stushnoff C. Downregulating alpha-galactosidase enhances freezing tolerance in transgenic petunia. *Plant Physiol* 2003; 133:901-9.
46. Peters S, Mundree SG, Thomson JA, Farrant JM, Keller F. Protection mechanisms in the resurrection plant *Xerophyta viscosa* (Baker): both sucrose and raffinose family oligosaccharides (RFOs) accumulate in leaves in response to water deficit. *J Exp Botany* 2007; 58:1947-56.
47. McPhee KE, Zematra RS, Brown J, Myers JR. Genetic analysis of the raffinose family oligosaccharides in common bean. *J Amer Hort Sci* 2002; 127:376-82.
48. Rimbach G, Pallauf J, Moehring J, Kraemer K, Minihane AM. Effect of dietary phytate and microbial phytase on mineral and trace element bioavailability—A literature review. *Curr Topics in Nutraceut Res* 2008; 6:131-44.
49. Foley RC, Singh KB. Isolation of a *Vicia faba* metallothionein-like gene: expression in foliar trichomes. *Plant Mol Biol* 1994; 26:435-44.
50. Foley RC, Liang ZM, Singh KB. Analysis of type 1 metallothionein cDNAs in *Vicia faba*. *Plant Mol Biol* 1997; 33:583-91.
51. Cobbett C, Goldsbrough P. Phytochelatin and metallothioneins: Roles in heavy metal detoxification and homeostasis. *Ann Rev Plant Biol* 2002; 53:159-82.
52. Guo WJ, Meetam M, Goldsbrough PB. Examining the specific contributions of individual Arabidopsis metallothioneins to copper distribution and metal tolerance. *Plant Physiol* 2008; 146:1697-706.
53. Leszczyszyn OI, Schmid R, Blindauer CA. Toward a property/function relationship for metallothioneins: histidine coordination and unusual cluster composition in a zinc-metallothionein from plants. *Proteins* 2007; 68:922-35.
54. Kägi JHR, Schäffer A. Biochemistry of metallothionein. *Biochem* 1988; 27:8509-15.
55. Chimenti F, Aouffen M, Favier A, Seve M. Zinc homeostasis-regulating proteins: new drug targets for triggering cell fate. *Curr Drug Targets* 2003; 4:323-38.
56. Davis SR, Cousins RJ. Metallothionein expression in animals: a physiological perspective on function. *J Nutr* 2000; 130:1085-8.
57. Kamel HA. Lead Accumulation and its Effect on Photosynthesis and Free Amino Acids in *Vicia faba* Grown Hydroponically. *Austr J Basic and Appl Sci* 2008; 2:438-46.
58. Piechalak A, Tomaszewska B, Baralkiewicz D, Malecka A. Accumulation and detoxification of lead ions in legumes. *Phytochem* 2002; 60:153-62.
59. Srivastava S, Mishra S, Dwivedi S, Baghel VS, Verma S, Tandon PK, et al. Nickel phytoremediation potential of broad bean, *Vicia faba* L., and its biochemical responses. *Bull Environ Contam Toxicol* 2005; 74:715-24.
60. Bond DA, Jellis GJ, Rowland GG, Le Guen J, Robertson LD, Khalil SA, Ji-Juan L. Present status and future strategy in breeding faba beans (*Vicia faba* L.) for resistance to biotic and abiotic stresses. *Euphytica* 1994; 73:151-66.
61. Zhang Y, Lewis K. Fabatins: new antimicrobial plant peptides. *FEMS Microbiol Lett* 1997; 149:59-64.
62. Stiekema WJ, Heidekamp F, Dirkse WG, van Beckum J, de Haan P, ten Bosch C, Louwse JD. Molecular cloning and analysis of four potato tuber mRNAs. *Plant Mol Biol* 1988; 11:255-69.
63. Silverstein KA, Moskal WA Jr, Wu HC, Underwood BA, Graham MA, Town CD, VandenBosch KA. Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. *Plant J* 2007; 51:262-80.
64. Bahgat S, Shabban OA, El-Shihy O, Lightfoot DA, El-Shemy HA. Establishment of the Regeneration System for *Vicia faba* L. *Curr Issues Mol Biol* 2009; 11:47-54.