



# Transcriptome profiling identifies candidate genes associated with the accumulation of distinct sulfur $\gamma$ -glutamyl dipeptides in *Phaseolus vulgaris* and *Vigna mungo* seeds

Dengqun Liao<sup>1</sup>, Dustin Cram<sup>2</sup>, Andrew G. Sharpe<sup>2</sup> and Frédéric Marsolais<sup>1,3\*</sup>

<sup>1</sup> Genomics and Biotechnology, Southern Crop Protection and Food Research Centre, Agriculture and Agri-Food Canada, London, ON, Canada

<sup>2</sup> National Research Council Canada, Saskatoon, SK, Canada

<sup>3</sup> Department of Biology, University of Western Ontario, London, ON, Canada

## Edited by:

Scott Jackson, University of Georgia, USA

## Reviewed by:

Dongying Gao, University of Georgia, USA

Yupeng Li, University of Georgia, USA

## \*Correspondence:

Frédéric Marsolais, Genomics and Biotechnology, Southern Crop Protection and Food Research Centre, Agriculture and Agri-Food Canada, 1391 Sandford Street, London, ON N5V 4T3, Canada.  
e-mail: frederic.marsolais@agr.gc.ca

Common bean (*Phaseolus vulgaris*) and black gram (*Vigna mungo*) accumulate  $\gamma$ -glutamyl-S-methylcysteine and  $\gamma$ -glutamyl-methionine in seed, respectively. Transcripts were profiled by 454 pyrosequencing data at a similar developmental stage coinciding with the beginning of the accumulation of these metabolites. Expressed sequence tags were assembled into Unigenes, which were assigned to specific genes in the early release chromosomal assembly of the *P. vulgaris* genome. Genes involved in multiple sulfur metabolic processes were expressed in both species. Expression of *Sultr3* members was predominant in *P. vulgaris*, whereas expression of *Sultr5* members predominated in *V. mungo*. Expression of the cytosolic SERAT1;1 and -1;2 was approximately fourfold higher in *P. vulgaris* while expression of the plastidic SERAT2;1 was twofold higher in *V. mungo*. Among BSAS family members, *BSAS4;1*, encoding a cytosolic cysteine desulfhydrase, and *BSAS1;1*, encoding a cytosolic O-acetylserine sulphydrylase were most highly expressed in both species. This was followed by *BSAS3;1* encoding a plastidic  $\beta$ -cyanoalanine synthase which was more highly expressed by 10-fold in *P. vulgaris*. The data identify *BSAS3;1* as a candidate enzyme for the biosynthesis of S-methylcysteine through the use of methanethiol as substrate instead of cyanide. Expression of *GLC1* would provide a complete sequence leading to the biosynthesis of  $\gamma$ -glutamyl-S-methylcysteine in plastids. The detection of S-methylhomogluthathione in *P. vulgaris* suggested that homogluthathione synthetase may accept, to some extent,  $\gamma$ -glutamyl-S-methylcysteine as substrate, which might lead to the formation of S-methylated phytochelatins. In conclusion, 454 sequencing was effective at revealing differences in the expression of sulfur metabolic genes, providing information on candidate genes for the biosynthesis of distinct sulfur amino acid  $\gamma$ -glutamyl dipeptides between *P. vulgaris* and *V. mungo*.

**Keywords:** sulfur metabolism, gamma-glutamyl dipeptides, S-methylcysteine, *Vigna mungo*, *Phaseolus vulgaris*, developing seed, 454 transcriptome sequencing

## INTRODUCTION

Among grain legumes, common bean (dry bean, *Phaseolus vulgaris* L.) is dominant for direct human consumption. Like other grain legumes, its protein quality is limited by the sub-optimal levels of the essential sulfur amino acids, Met, and Cys. Common bean accumulates high levels of a non-protein sulfur amino acid in seed, S-methyl-Cys, partly in the form of a  $\gamma$ -Glu dipeptide (Giada et al., 1998; Taylor et al., 2008).  $\gamma$ -Glu-S-methyl-Cys accumulates exclusively in seed and not in other tissues (Watanabe et al., 1971). The non-protein amino acid, S-methyl-Cys, cannot substitute for Met or Cys in the diet (Padovese et al., 2001). The taxonomic distribution of S-methyl-Cys encompasses all *Phaseolus* and several *Vigna* species, notably cowpea (*Vigna unguiculata*), and mung bean (*Vigna radiata*), but excludes black gram (*Vigna mungo*) and adzuki bean (*Vigna angularis*), two Asian *Vigna* species (Baldi and Salamini, 1973; Evans and Boulter, 1975;

Kasai and Larsen, 1980; Kasai et al., 1986). *V. mungo* accumulates the dipeptide  $\gamma$ -Glu-Met rather than  $\gamma$ -Glu-S-methyl-Cys (Otoul et al., 1975). While S-methyl-Cys may be formed by direct methylation of Cys in radish, this does not appear to be the case in common bean (Thompson and Gering, 1966). Rather, it was proposed to be formed by the condensation of O-acetylserine and methanethiol, catalyzed by O-acetylserine sulphydrylase (also designated as Cys synthase), a member of the  $\beta$ -substituted Ala synthase (BSAS) family of enzymes. O-Acetylserine sulphydrylase partially purified from common bean seedlings catalyzes this reaction *in vitro* (Smith and Thompson, 1971). This reaction also takes place in *Arabidopsis* cell cultures grown in the presence of a high concentration of Met (Rébeillé et al., 2006). The detection of small amounts of S-methylhomogluthathione in *V. radiata* seeds suggests that homogluthathione synthetase may accept to some extent  $\gamma$ -Glu-S-methyl-Cys as substrate (Kasai et al., 1986).

In common bean, a deficiency of the major seed storage proteins, the 7S globulin phaseolin, and the lectin phytohemagglutinin results in a shift of sulfur from *S*-methyl-Cys to the Cys and Met pools in protein (Taylor et al., 2008). Lack of the major storage proteins is compensated by increased levels of several sulfur-rich proteins, including the 11S globulin legumin, albumin-2, defensin D1, albumin-1, and Bowman–Birk type proteinase inhibitor (Marsolais et al., 2010; Yin et al., 2011). Analysis of microarray data revealed increased levels of transcripts for two additional types of sulfur-rich proteins, the basic 7S globulin and Kunitz trypsin inhibitor, and a general up-regulation of transcripts coding for sulfate transporters and enzymes of sulfur amino acid metabolism (Liao et al., 2012). Up-regulation of *SERAT1;1* and *-1;2* expression indicated an activation of cytosolic *O*-acetylserine biosynthesis, while down-regulation of *SERAT2;1* encoding a predicted plastidic enzyme suggested that Cys and *S*-methyl-Cys biosynthesis may take place in different subcellular compartments, providing an explanation for their opposite regulation.

In the present study, Roche 454 Genome Sequencer FLX pyrosequencing was utilized to profile the developing seed transcriptome of the two closely related species, *P. vulgaris* and *V. mungo*, both members of subtribe Phaseolinae, which differ by the accumulation of  $\gamma$ -Glu-*S*-methyl-Cys and  $\gamma$ -Glu-Met, respectively. In *P. vulgaris*, this technology has been used to investigate transcripts from leaf, flower, root, and pod (Kalavacharla et al., 2011; Liu et al., 2012), whereas Illumina's sequencing by synthesis technology has been used to profile the seed transcriptome (Young et al., 2011; Zhai et al., 2011) and identify microRNAs (Peláez et al., 2012). Transcript profiling data generated by high-throughput sequencing has not been reported yet from *V. mungo* but is available from *V. radiata* and *V. unguiculata* (Barrera-Figueroa et al., 2011; Kido et al., 2011; Moe et al., 2011). Differences in the expression of sulfate transporters and enzymes of sulfur amino acid metabolism suggest adaptations which may be related to the accumulation of distinct sulfur amino acid  $\gamma$ -Glu dipeptides in developing seed.

## MATERIALS AND METHODS

### PLANT MATERIALS AND GROWTH CONDITIONS

The common bean (*P. vulgaris*) line BAT93 (Nodari et al., 1993) and black gram (*V. mungo*) cultivar Barimash-2 (Afzal et al., 2002) were grown in growth cabinets (Convicon E8H, Controlled Environments, Winnipeg, MB, Canada). Seeds were sown in small pots (8 cm  $\times$  12 cm) containing Pro-Mix PGX (Premier Horticulture, Rivière-du-Loup, QC, Canada). Seedlings were transplanted to larger pots (17 cm  $\times$  20 cm) containing Pro-Mix BX (Premier Horticulture). Plants were given 16 h light (300–400  $\mu$ mol photons  $m^{-2} s^{-1}$ ) and 8 h dark, with a temperature cycling between 18°C and 24°C (Pandurangan et al., 2012). Plants were fertilized with 20:20:20 (Plant-Prod, Laval, QC, Canada) with the equivalent of 0.1 g per pot for Barimash-2 and 0.3 g per pot for BAT93 after transplantation and subsequently every 2 weeks. Developing seeds were harvested, immediately frozen in liquid nitrogen, and stored at -80°C.

### EXTRACTION AND QUANTIFICATION OF FREE AMINO ACIDS

Amino acids were extracted from seeds and quantified by HPLC after derivatization with phenylisothiocyanate as previously

described (Yin et al., 2011). For *V. mungo*, the solvent used for extraction, ethanol:water (70:30) was supplemented with 2% (w/v) polyvinylpyrrolidone to neutralize phenolic compounds.  $\gamma$ -Glu-Leu,  $\gamma$ -Glu-Met, and *S*-methylhomoglutathione standards were from Bachem Americas (Torrance, CA, USA). Free amino acid profiles were analyzed with CLUSTER 3.0 (de Hoon et al., 2004).

### RNA EXTRACTION, cDNA LIBRARY PREPARATION, AND SEQUENCING

RNA was extracted from developing seeds according to Wang and Vodkin (1994). RNA was quantified by spectrophotometry with a NanoDrop 1000 (Thermo Scientific, Wilmington, DE, USA) and its quality evaluated from  $A_{260/280}$  ratio and by agarose gel electrophoresis. Total RNA was treated with amplification-grade DNase I (Life Technologies, Burlington, ON, Canada) and extracted with acidic phenol:chloroform. For purification with the Poly(A) Purist MAG Kit, 480  $\mu$ g of DNaseI treated total RNA was used, and this was followed by treatment with the RiboMinus Plant Kit for RNA-Seq (Life Technologies). For cDNA Rapid Library Prep, 400 ng of the purified samples was used. Three copies per bead were used for emulsion-based clonal amplification and  $2 \times 10^6$  beads were loaded per half-plate for 454 pyrosequencing on a Genome Sequencer-FLX Instrument using Titanium chemistry (Roche Applied Science, Laval, QC, Canada).

### SEQUENCE ANALYSIS

The 454 read pre-processing steps including filtering based on read quality score, adapter trimming, repeat filtering, low-complexity masking, and poly(A/T) removal were done with custom, in-house scripts of the Bioinformatics lab at National Research Council Canada in Saskatoon, SK, Canada. Assembly of the cleaned 454 ESTs was performed with iAssembler v1.3.0 using default parameters: a maximum of 30bp long end clips, a minimum of 40bp overlap, and 97% identity (Zheng et al., 2011). Stand-alone BLAST (2.2.27+) from NCBI was used with a threshold *e*-value  $< 1 \times 10^{-10}$  (Park et al., 2012). Gene annotation was performed by BLASTX against *Arabidopsis thaliana* TAIR10 proteins (TAIR10\_pep\_20101214) (Lamesch et al., 2012). The TAIR ID with the top hit was extracted and used for functional annotation and gene ontology classification (Berardini et al., 2004). ESTs were searched by BLASTX against *Glycine max* proteins from the Glyma1 chromosome-based assembly (Gmax\_109\_peptide.fa)<sup>1</sup> (Schmutz et al., 2010). Unigenes were associated with specific genes in the *P. vulgaris* genome by BLASTX and BLASTN against the amino acid and gene coding sequences from the early release genome of the common bean (Pvulgaris\_v1.0\_218\_cds.fa) (Goodstein et al., 2012). These sequence data were produced by the United States Department of Energy Joint Genome Institute<sup>2</sup> in collaboration with United States Department of Agriculture–National Institute of Food and Agriculture. *P. vulgaris* Unigenes were assigned to a gene in the *P. vulgaris* genome on the basis of the top BLASTN hit and sequences were at least 97% identical. *V. mungo* Unigenes were assigned to a *P. vulgaris* gene in the same manner and sequences were at least 88% identical. BLASTX results

<sup>1</sup>[http://www.phytozome.net/commonbean\\_er](http://www.phytozome.net/commonbean_er)

<sup>2</sup><http://www.jgi.doe.gov/>

were used to confirm the assignments. When multiple Unigenes were assigned to a single gene, the number of reads associated with each of the Unigenes was summed to provide the total number of reads for this gene. When multiple Unigenes were assigned to a single gene, their alignment was usually non-overlapping, i.e., there were gaps in coverage of the reference coding sequence. A nomenclature was developed for sulfate transporters and sulfur metabolic genes on the basis of phylogenetic clustering with *Arabidopsis* and soybean homologs (Yi et al., 2010; Hell and Wirtz, 2011; Takahashi et al., 2011; Liao et al., 2012).

#### ACCESSION NUMBERS

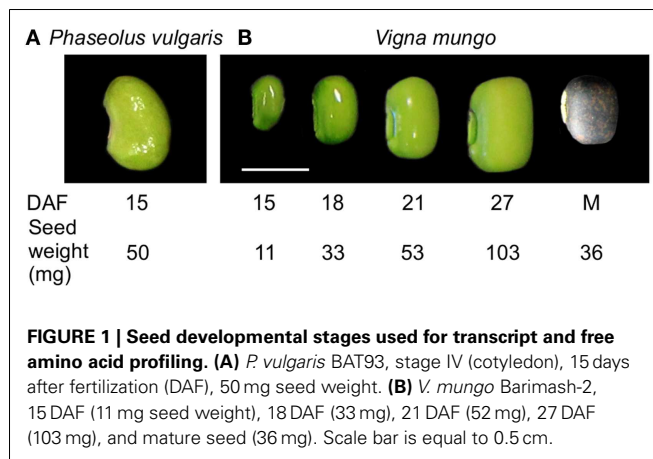
Sequencing data are available at SRA<sup>3</sup> under the accession numbers SRS360251 for *P. vulgaris* and SRS360666 for *V. mungo*. The *P. vulgaris* data is composed of two subsets: GIWE0SV (accession number SRR553466) and GMAPAYS (accession number SRR553467). Only the GMAPAYS subset was used for assembly and data analysis as the GIWE0SV dataset was the result of an additional trial run.

## RESULTS

### METABOLIC PROFILING OF FREE AMINO ACIDS DURING SEED DEVELOPMENT IN *V. MUNGO*

Stage IV, corresponding to the cotyledon stage, with an average seed weight of 50 mg, at 15 days after flowering (DAF) (Figure 1) (Walbot et al., 1972) was selected for transcript profiling of developing *P. vulgaris* seeds because it corresponds to the beginning of the active phase of  $\gamma$ -Glu-S-methyl-Cys accumulation, before free S-methyl-Cys levels start to decline (Yin et al., 2011). To identify a physiologically similar stage in *V. mungo*, free amino acids, including the dipeptides  $\gamma$ -Glu-Leu and  $\gamma$ -Glu-Met (Otoul et al., 1975), were extracted and profiled at five stages of seed development (Figure 1). Figure 2A shows that Asn, followed by Gln, and Ala, were the dominant free amino acids at early stages of development, while  $\gamma$ -Glu-Met was the most abundant at maturity. Amino acid concentration was normalized over the average of developmental stages, expressed in a  $\log_2$  scale, and *k*-means analysis performed to reveal common developmental patterns. Amino acids were grouped into six clusters (Figure 2B). The levels of most amino acids, present in the first four clusters, were stable from 15 to 21 DAF and declined abruptly at 27 DAF suggesting their incorporation into storage proteins. The fifth cluster grouped amino acids with levels similar or greater than average at 27 DAF and included Phe, Tyr, Orn, as well as  $\gamma$ -Glu-Leu, whose levels increased steadily except between 18 and 21 DAF.  $\gamma$ -Glu-Met was clustered separately from other amino acids because its levels increased at all developmental stages. Eighteen DAF was selected for transcript profiling as it corresponded to the beginning of the active phase of  $\gamma$ -Glu-Met accumulation when free Met was at its highest levels. In *P. vulgaris*, storage protein transcripts are most abundant beginning at stage IV, which precedes storage product accumulation (Bobb et al., 1995), whereas phases of cell expansion and synthesis of reserve metabolites more completely overlap in *V. mungo* (Sehgal and Gandhi, 1987).

<sup>3</sup><http://www.ncbi.nlm.nih.gov/sra/>



### QUANTIFICATION OF S-METHYLHOMOGLUTATHIONE IN *P. VULGARIS* SEEDS

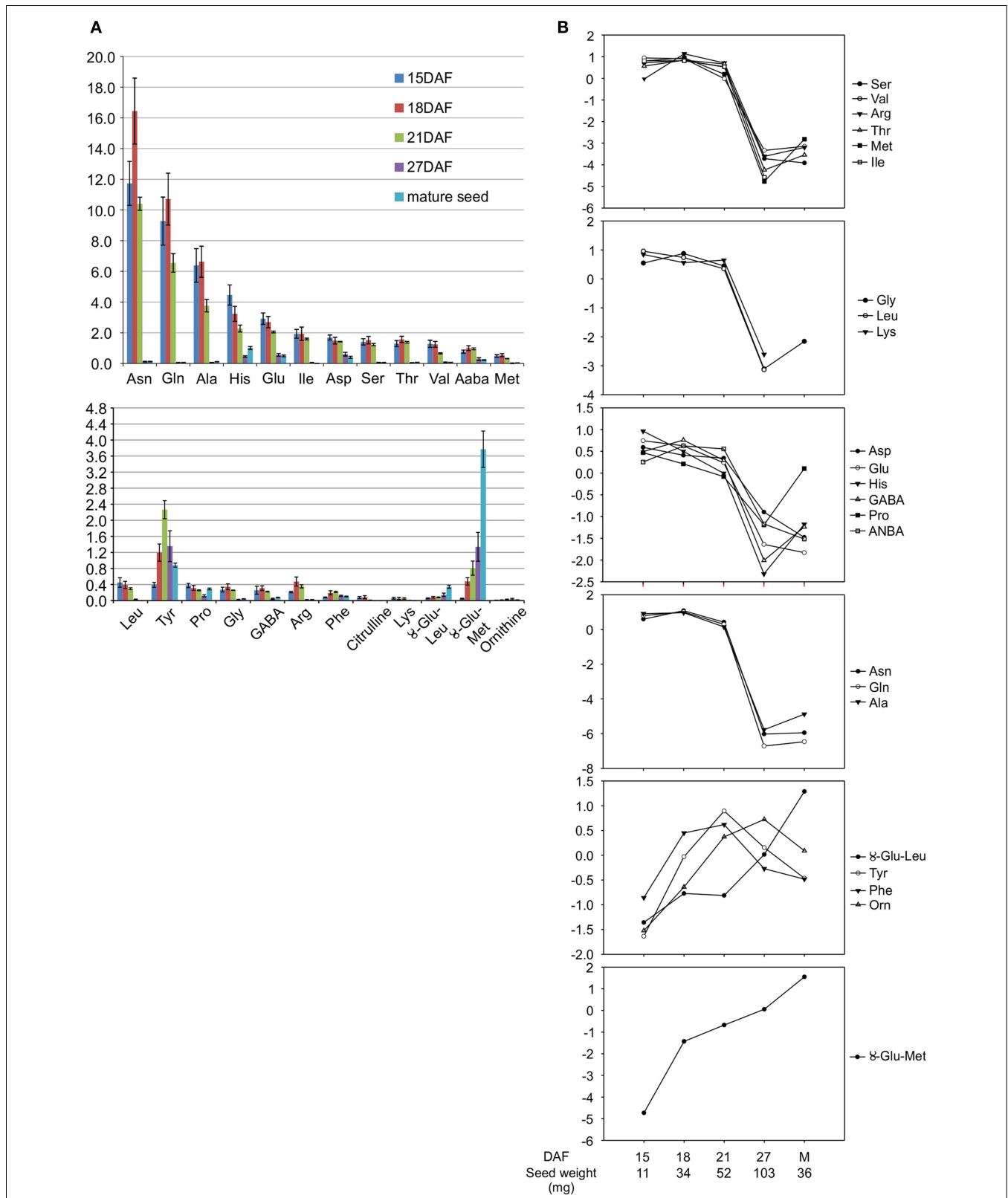
To determine whether S-methylhomoglutathione accumulates in *P. vulgaris* as in *V. radiata* (Kasai et al., 1986), free amino acids were extracted from mature seeds of the BAT93 line. S-Methylhomoglutathione was quantified in comparison with a synthetic standard. The concentration measured was equal to  $73.3 \pm 6.8$  pmol per mg seed weight (average  $\pm$  standard deviation;  $n = 3$ ), and is similar to that determined in *V. radiata* seeds (Kasai et al., 1986).

### TRANSCRIPT PROFILING OF DEVELOPING SEEDS FROM *P. VULGARIS* AND *V. MUNGO* BY 454 PYROSEQUENCING

RNA was extracted from developing seeds of *P. vulgaris* and *V. mungo* at the selected developmental stages. Transcripts were profiled on a 454 Genome Sequencer-FLX Titanium platform. A half-plate was sequenced for each sample starting from the same amount of input cDNA. Table 1 lists general statistics for the 454 transcriptome data. The number of expressed sequenced tags (ESTs) after pre-processing was remarkably similar between the two species, within a 0.3% difference. *V. mungo* ESTs were slightly longer on average by 9%. The number of assembled contigs was very close between the two species, within a 5% difference. The number of singletons was considerably higher in *P. vulgaris*, by 38%, which correlated with the shorter EST length. Importantly, the number of Unigenes with significant BLAST hits to the *Arabidopsis*, soybean, and common bean genomes was almost equal between the two species. An analysis of gene ontology categories showed a comparable distribution between the two samples (Figure 3). Overall, these data suggested a similar physiological state between samples. These data validated the selection of developmental stages from the two species, and supported further comparative analyzes of the expression of sulfur metabolic genes.

### EXPRESSION OF SULFATE TRANSPORTER AND SULFUR METABOLIC TRANSCRIPTS

To investigate possible differences in expression of sulfate transporter or sulfur metabolic genes, which may correlate with the accumulation of distinct sulfur amino acid  $\gamma$ -Glu dipeptides, Unigenes, and their corresponding ESTs were assigned to specific genes in the early release chromosomal assembly of the *P. vulgaris*



**FIGURE 2 | Free amino acid profiles during seed development in *V. mungo*. (A)** Free amino acid concentration expressed in nmol per mg seed weight;  $n = 3$ . Error bars represent standard deviation **(B)** *k*-Means cluster

analysis of free amino acid profiles during seed development. Data on the y axis is the amino acid concentration in **(A)** normalized to the average of developmental stages and expressed in a log<sub>2</sub> scale.

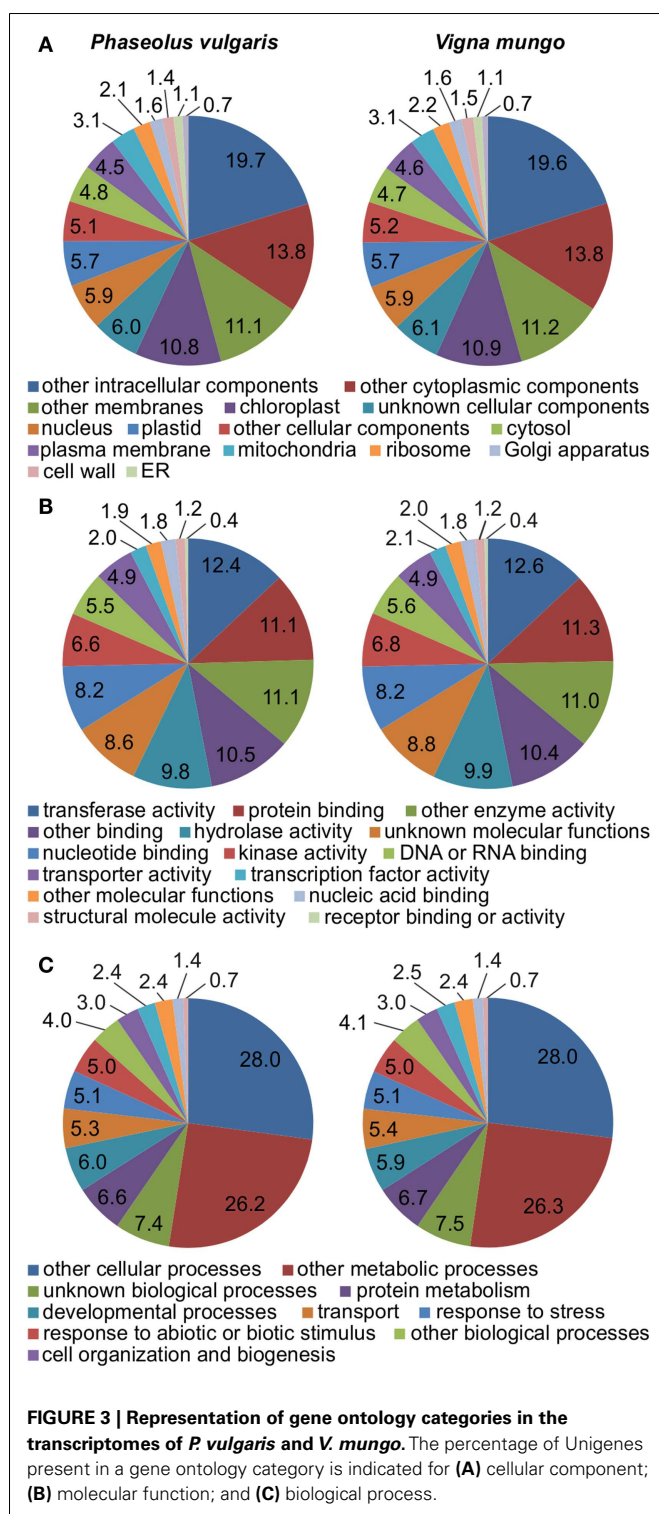


**Table 1 | Summary of 454 transcriptome data from *P. vulgaris* and *V. mungo* developing seeds.**

	<i>P. vulgaris</i>	<i>V. mungo</i>
Number of reads	757719	717036
Number of ESTs after pre-processing	662417	660318
Median length (bp)	358	408
Average length (bp)	339	369
Longest length (bp)	770	665
Mean GC content (%)	45.1	44.7
Number of contigs	38344	40125
Number of singletons	39104	28349
Number of Unigenes	77448	68474
Contig average length (bp)	751	809
Unigene average length (bp)	510	583
Longest contig (bp)	7845	6749
Number of Unigenes with BLASTX hit to TAIR10	40129	41430
Number of Unigenes with BLASTX hit to Glyma1	48524	48571
Number of Unigenes with BLASTX hit to <i>P. vulgaris</i> v1.0	51119	50165

Expressed sequence tags (ESTs) were assembled with *iAssembler* v1.3.0. BLAST hits were considered significant at an *e*-value  $< 1 \times 10^{-10}$ .

genome. A previously developed nomenclature based on phylogenetic clustering with *Arabidopsis* and soybean homologs was utilized (Table A1 in Appendix) (Liao et al., 2012). Sulfate transporter and sulfur metabolic genes are often members of multigene families and the corresponding gene products may have different subcellular localizations. Table 2 compares the numbers of genes in the *P. vulgaris*, *G. max*, and *Arabidopsis* genomes coding for sulfate transporters and the different classes of sulfur metabolic enzymes. The higher number of genes in *G. max* reflects its paleopolyploid nature as compared with the true diploids *P. vulgaris* and *Arabidopsis*. As compared with *Arabidopsis*, *P. vulgaris* has more genes coding for sulfate transporters and enzymes of glutathione/homoglutathione biosynthesis, but fewer genes involved in sulfate assimilation or activation (Table A1 in Appendix). Figure 4 presents the number of ESTs assigned to each gene in the sulfur metabolic pathway. This figure shows that genes involved in multiple sulfur metabolic processes, including sulfate transport, sulfate activation, sulfate assimilation, *de novo* Cys and Met biosynthesis, homoglutathione biosynthesis and Met catabolism were expressed in both species. Transcripts for 12 different sulfate transporter genes were detected in the *P. vulgaris* and *V. mungo* samples, out of a total of 19 encoded by the *P. vulgaris* genome. Transcripts of members of the *Sultr1* and *-2* families, encoding high and low affinity sulfate transporters involved in sulfate uptake by the root, and in the translocation of sulfate from root to shoot, respectively (Takahashi et al., 2000; Yoshimoto et al., 2007; Barberon et al., 2008), were detected at very low levels. Rather, group 3 genes were the most highly expressed in *P. vulgaris*, with 95 ESTs representing five different genes, as compared with 32 ESTs assigned to the same genes in *V. mungo*. Group 3 genes have been associated with the transfer of sulfate from the seed coat to the embryo in *Arabidopsis* (Zuber et al., 2010). Transcripts of group 5 genes were dominant in *V. mungo*, with 106 ESTs for four different genes, as compared



**FIGURE 3 | Representation of gene ontology categories in the transcriptomes of *P. vulgaris* and *V. mungo*.** The percentage of Unigenes present in a gene ontology category is indicated for (A) cellular component; (B) molecular function; and (C) biological process.

with 11 ESTs for the same genes in *P. vulgaris*. *Sultr5;2* has been implicated as a high affinity molybdate transporter in *Arabidopsis* (Tomatsu et al., 2007; Baxter et al., 2008), although expression of the corresponding gene in wheat under sulfur deficiency did not correlate with molybdenum accumulation (Shinmachi et al., 2010). The present data suggest that at least some of the *Sultr5* members may function as sulfate transporters in *V. mungo* seeds.

**Table 2 | Number of genes coding for sulfate transporters and sulfur metabolic enzymes in the genomes of *P. vulgaris*, *G. max*, and *Arabidopsis*.**

	<i>P. vulgaris</i>	<i>G. max</i>	<i>Arabidopsis</i>
Sulfate transporter	19	35	14
Sulfate adenylyl- transferase	2	4	4
Adenylyl-sulfate kinase	3	5	4
Adenylyl-sulfate reductase	2	3	3
Sulfite reductase	1	2	1
Sulfite oxidase	1	2	1
Serine acetyltransferase	6	9	5
$\beta$ -Substituted Alasynthase	8	18	9
Cystathionine $\gamma$ -synthase	1	2	2
Cystathionine $\beta$ -lyase	1	2	1
Homocysteine S-methyltransferase	3	4	3
Metsynthase	3	7	3
Met $\gamma$ -lyase	2	3	1
Glu-Cys ligase	2	4	1
Glutathione synthetase	3	3	1

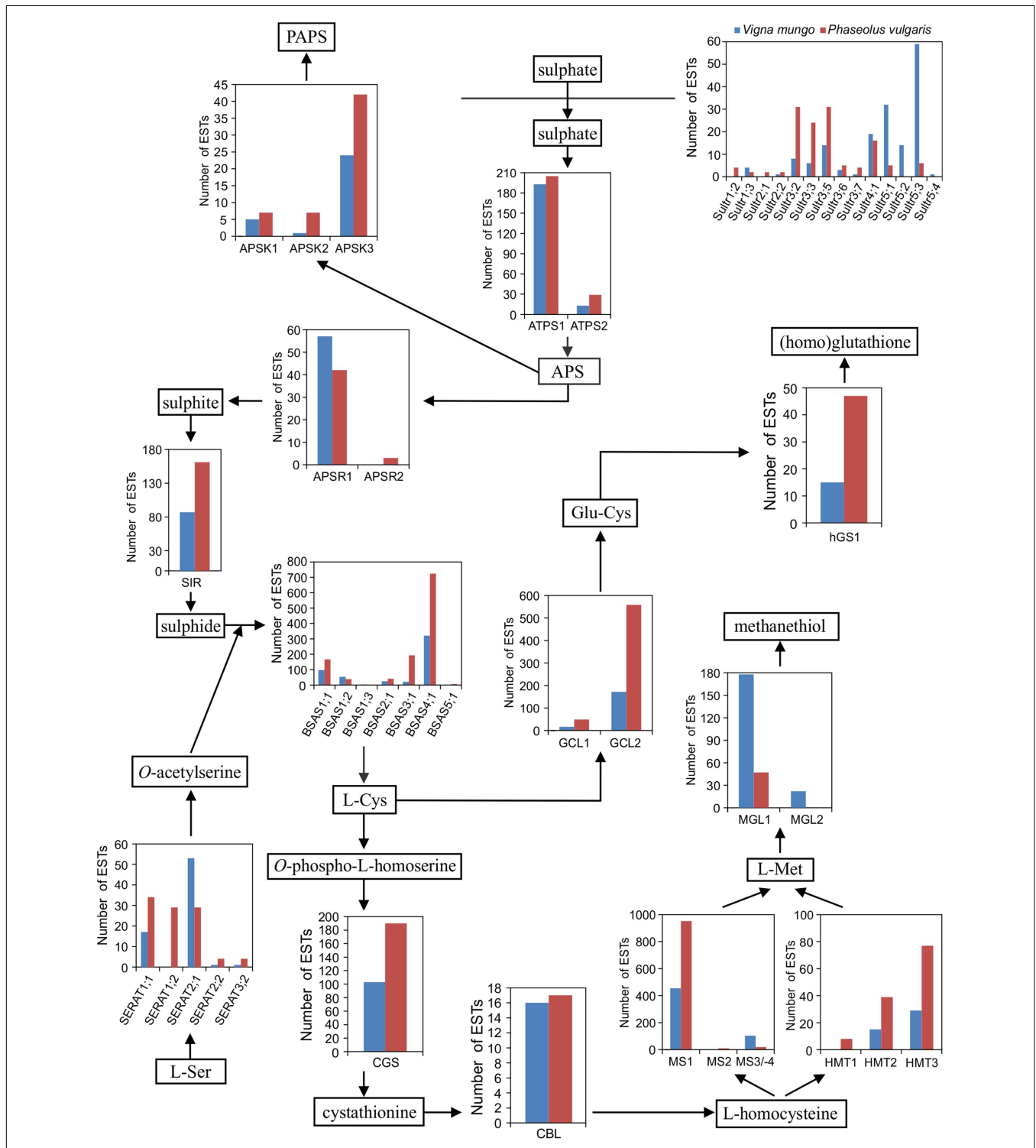
The first step in sulfate assimilation is catalyzed by sulfate adenylyltransferase (ATPS). Its product, adenosine 5'-phosphosulphate, is found at a branch point between sulfate activation and reductive sulfate assimilation (Mugford et al., 2009, 2011; Kopriva et al., 2012). Among the two ATPS genes, *ATPS1* was predominantly expressed in both species. Adenylyl-sulfate kinase (APSK) is responsible for the biosynthesis of the universal sulfate donor, 3'-phosphoadenosine 5'-phosphosulphate. Of the three APSK genes, *APSK3* was the most highly expressed. In the other branch of the pathway leading to reductive sulfate assimilation, adenylyl-sulfate reductase (APSR) produces sulfite from adenosine 5'-phosphosulphate. *APSR1* was the main isoform expressed in seed. *APSR2* was expressed at lower levels, by more than 10-fold and detected only in *P. vulgaris*. Sulfite reductase (SIR) is encoded by a single gene in *P. vulgaris* and its expression was detected in both species. The same was observed for sulfite oxidase, with 14 ESTs in *P. vulgaris* as compared with 27 in *V. mungo*. Overall, the genes involved in sulfate assimilation and activation were expressed at relatively similar levels between *P. vulgaris* and *V. mungo*.

Sulfide is combined with *O*-acetylserine produced by serine acetyltransferase (SERAT) to form Cys, in a reaction catalyzed by *O*-acetylserine sulphydrylase. *P. vulgaris* had a higher number of ESTs for *SERAT1*;1 and -1;2 than *V. mungo* by close to fourfold, while transcripts for *SERAT2*;1 were more abundant in *V. mungo* by approximately twofold. According to phylogeny and analysis with WoLF PSORT (Horton et al., 2007), *SERAT1*;1 and -1;2 are predicted to be localized in the cytosol and *SERAT2*;1 in plastids. *Arabidopsis* has two *SERAT2* isoforms localized in plastids and mitochondria, respectively (Noji et al., 1998), and a fusion between soybean *SERAT2*;1 and a fluorescent protein was detected in plastids and the cytosol (Liu et al., 2006). Among BSAS family members, *BSAS4*;1 was the most abundantly expressed followed by *BSAS1*;1. The number of ESTs associated with these two genes was approximately twofold higher in *P. vulgaris* than in *V.*

*mungo*. *BSAS1*;1 encodes a cytosolic *O*-acetylserine sulphydrylase in both soybean and *Arabidopsis* (Heeg et al., 2008; Watanabe et al., 2008; Hell and Wirtz, 2011; Yi and Jez, 2012; Yi et al., 2012). *BSAS4*;1 clusters with *GmBSAS4*;2, encoding a cytosolic BSAS which preferentially catalyzes the desulfhydration of Cys (*GmDES*; Glyma10g30140) (Yi et al., 2010; Yi and Jez, 2012). *Arabidopsis* *BSAS4*;3 also represents a Cys desulfhydrase (Alvarez et al., 2010). *BSAS3*;1 was also abundantly expressed in *P. vulgaris*, with a number of ESTs approximately ninefold higher than in *V. mungo*. In *Arabidopsis*, *BSAS3*;1 is a  $\beta$ -cyanoalanine synthase involved in the detoxication of cyanide. This mitochondrial enzyme catalyzes the formation of  $\beta$ -cyanoalanine from Cys and cyanide releasing sulfide (Watanabe et al., 2008). Likewise, soybean *BSAS3*;1 is a  $\beta$ -cyanoalanine synthase (Yi and Jez, 2012; Yi et al., 2012). Interestingly, common bean *BSAS3*;1 is predicted by WoLF PSORT to be localized in plastids rather than mitochondria.

In the steps leading to the formation of Met, Cys is condensed by cystathionine  $\gamma$ -synthase (CGS) with *O*-phospho-L-homoserine to form cystathionine. Cystathionine is then cleaved by cystathionine  $\beta$ -lyase (CBL) producing homocysteine which can be utilized for Met biosynthesis. Both steps catalyzed by CGS and CBL occur in plastids. Homocysteine can be methylated by homocysteine S-methyltransferase using S-methyl-Met as methyl donor, a major form of organic sulfur transported to seed (Bourgis et al., 1999; Ranocha et al., 2000; Lee et al., 2008; Tan et al., 2010). Alternatively, homocysteine can be methylated by Met synthase (MS) using methyltetrahydrofolate as methyl donor. MS is present in both plastids and the cytosol. However, it is assumed that only plastidic MS has a significant contribution to *de novo* Met biosynthesis, while cytosolic MS is involved in the recycling of Met as part of the S-adenosyl-Met cycle (Amir and Hacham, 2008; Jander and Joshi, 2010). Both CGS and CBL are encoded by unique genes in *P. vulgaris*. The number of ESTs was almost twofold higher for CGS in *P. vulgaris* than in *V. mungo*, but similar for CBL. Among the three HMT genes, *HMT3* was predominantly expressed in both species, followed by *HMT2* and -1. The number of ESTs was higher in *P. vulgaris* for all three genes, on average by almost threefold. Among MS genes, *MS1*, encoding a predicted cytosolic enzyme, was the most highly expressed, with a twofold higher number of ESTs in *P. vulgaris* than in *V. mungo*. ESTs corresponding to the plastidic MS were more abundant in *V. mungo*, by approximately fivefold. In the early release chromosomal assembly of the *P. vulgaris* genome, the gene coding for plastidic MS is split into two, *MS3*, encoding an N-terminal segment and *MS4*, encoding a C-terminal segment, located on chromosomes 5 and 6, respectively. The present EST data do not support this split since a contig overlapped between the two segments, which suggests a possible artifact from the genome assembly process. The enzyme Met  $\gamma$ -lyase catalyzes the breakdown of Met, producing methanethiol, ammonia, and 2-oxobutanoate, which can be utilized for Ile biosynthesis (Rébeillé et al., 2006; Goyer et al., 2007; Joshi and Jander, 2009). There are two MGL genes in the *P. vulgaris* genome. Only *MGL1* was expressed in *P. vulgaris*, while both genes were expressed in *V. mungo*. The number of ESTs for *MGL1* was approximately threefold higher in *V. mungo*.

Beside Met biosynthesis, Cys can be used to synthesize glutathione and homoglutathione (Galant et al., 2011). The enzyme



**FIGURE 4 | Expression of sulfate transporter and sulfur metabolic genes in developing seed of *P. vulgaris* and *V. mungo*.** ESTs were assigned to individual genes in the chromosomal assembly of the early release of the *P. vulgaris* genome. Abbreviations are as follows: Sultr: sulfate transporter; ATPS: sulfate adenylyltransferase; APS: adenosine 5'-phosphosulfate; APSK: adenylyl-sulfate kinase; PAPS:

3'-phosphoadenosine 5'-phosphosulfate; APSR: adenylyl-sulfate reductase; SIR: sulfite reductase; SERAT: Ser acetyltransferase; BSAS: β-substituted Ala synthase; CGS: cystathionine γ-synthase; CBL: cystathionine β-lyase; HMT: homocysteine S-methyltransferase; MS: Met synthase; MGL: Met γ-lyase; GCL: Glu-Cys ligase; hGS: homoglutathione synthetase.

$\gamma$ -Glu-Cys ligase (GCL) catalyzes the formation of  $\gamma$ -Glu-Cys. There are two *GCL* genes in the *P. vulgaris* genome. *GCL1* is predicted to be localized in the chloroplast and *GCL2* in the cytosol, in agreement with subcellular localization data from spinach and pea leaf, and common bean and cowpea nodule (Hell and Bergmann, 1990; Moran et al., 2000). *GCL2* was expressed at higher levels than *GCL1* in both species, by approximately 10-fold. Both *GCL1* and *-2* were expressed at higher levels in *P. vulgaris* than in *V. mungo*, by approximately threefold. The *P. vulgaris* genome contains a single gene encoding homoglutathione synthetase (*hGS1*) and two genes encoding glutathione synthetase (*GS1* and *-2*) located in tandem at a unique locus, like in other legumes (Frendo et al., 2001). Only *hGS1* transcripts were detected in seed, and not those of *GS1* and *-2*. This result is consistent with the fact that the concentration of homoglutathione is 80-fold higher than the concentration of glutathione in *P. vulgaris* seeds (Klapheck, 1988). The number of *hGS1* ESTs was approximately threefold higher in *P. vulgaris* than in *V. mungo*. *hGS1* protein is predicted to be localized in the cytosol, consistent with previous reports on the subcellular localization of hGS activity in *Phaseolus coccineus* leaves and nodules of common bean (Klapheck et al., 1988; Moran et al., 2000). Phytochelatin synthase (PCS) uses homoglutathione to synthesize phytochelatin. There are two PCS genes in the *P. vulgaris* genome located in tandem. Both were expressed at low levels, with 1 EST each for *PCS1* in both species, while four *PCS2* ESTs were detected exclusively in *P. vulgaris*.

## DISCUSSION

The transcriptome profiling data reflects the complexity of sulfur metabolism at the cellular level. Despite the close relationship between species, and the similarity in physiological, biochemical, and gene expression profiles between samples, key differences were noted in the number of ESTs assigned to different sulfur metabolic genes. These differences in gene expression may reflect adaptations related to the accumulation of distinct sulfur  $\gamma$ -Glu dipeptides. The data may provide information on potential candidate genes involved in their biosynthesis.

A prior study focusing on gene expression analysis between a pair of related *P. vulgaris* genotypes contrasted in the accumulation of Cys and *S*-methyl-Cys indicated that increased levels of Cys were associated with enhanced gene expression of cytosolic *SERAT1;1* and *-1;2*, while lower levels of *S*-methyl-Cys were associated with decreased transcript levels for plastidic *SERAT2;1* (Liao et al., 2012). This finding suggested a spatial separation of Cys and *S*-methyl-Cys biosynthesis in different subcellular compartments, the cytosol, and plastids, respectively. These data appeared consistent with the fact that legume seeds rely primarily on cytosolic *de novo* Cys biosynthesis. Indeed, overexpression of a cytosolic *O*-acetylserine sulfhydrylase in soybean seed led to increased Cys levels (Kim et al., 2012), whereas overexpression of a plastidic *SERAT* did not alter sulfur amino acid concentration in chickpea seed (Tabe et al., 2010). In agreement with these findings, *BSAS1;1*, encoding a cytosolic *O*-acetylserine sulfhydrylase, was among the most highly expressed members of the BSAS family. Surprisingly, the member with the highest expression, *BSAS4;1* encodes a cytosolic Cys desulfhydrase. The significance of its high expression in seed is not clear. The largest difference in

expression between species was observed for *BSAS3;1* encoding a  $\beta$ -cyanoalanine synthase predicted to be localized in plastids. Interestingly,  $\beta$ -cyanoalanine synthase purified from aerial tissues of *Lathyrus latifolius* and from leaves of *Spinacia oleracea* catalyzed the formation of *S*-methyl-Cys from Cys and methanethiol at a relatively low rate corresponding to 1–4% of  $\beta$ -cyanoalanine synthase activity (Ikegami et al., 1988a,b). This activity reached 13% for the  $\beta$ -cyanoalanine synthase purified from developing seeds of *Vicia angustifolia* (Ikegami et al., 1989). Taken together, the previous implication of plastidial *SERAT2;1*, the approximately 10-fold higher expression of *BSAS3;1* in *P. vulgaris* as compared with *V. mungo*, encoding a plastidic  $\beta$ -cyanoalanine synthase, and the fact that this enzyme may catalyze a similar reaction with methanethiol as with hydrogen cyanide, leading to the formation of *S*-methyl-Cys suggests that *BSAS3;1* may be responsible for the formation of *S*-methyl-Cys in *P. vulgaris* seed. Met  $\gamma$ -lyase may provide a source of methanethiol for *S*-methyl-Cys biosynthesis. Contrary to expectations, a higher number of ESTs for *MGL* genes were observed in *V. mungo* than in *P. vulgaris*.

The enzyme  $\gamma$ -GCL may be involved in the biosynthesis of  $\gamma$ -Glu-*S*-methyl-Cys. The mammalian enzyme has substantial *in vitro* activity with *S*-methyl-Cys as substrate (Rathbun, 1967; Sekura and Meister, 1977). The presence of a plastidic  $\gamma$ -GCL encoded by *GCL1* expressed in *P. vulgaris* seed would enable a complete pathway of  $\gamma$ -Glu-*S*-methyl-Cys biosynthesis in this organelle. The presence of a low concentration of *S*-methylhomoglutathione suggests that *hGS1* may accept, at least to a limited extent,  $\gamma$ -Glu-*S*-methyl-Cys as substrate. *In vitro* data indicates that *PCS1* can utilize *S*-alkylated glutathione for the synthesis of *S*-alkyl phytochelatin with a high efficiency (Vatamaniuk et al., 2000). This suggests that some *S*-methyl-Cys may accumulate in *S*-methylated phytochelatin, which may account, at least in part, for the discrepancy between concentrations measured for  $\gamma$ -Glu-*S*-methyl-Cys and total *S*-methyl-Cys in *P. vulgaris* seed (Taylor et al., 2008).

In summary, 454 transcriptome sequencing has been used to investigate the expression of sulfur metabolic genes in developing seeds of related *Phaseolinae* species accumulating a different sulfur  $\gamma$ -Glu dipeptide. Despite an overall similar physiological, biochemical, and gene expression profile between the two species, this technique was efficient at revealing differences in the expression of sulfur metabolic genes which provide information on candidate genes for the biosynthesis of these metabolites.

## ACKNOWLEDGMENTS

This research was supported in part by the Natural Products Genomics Resource (NAPGEN) consortium, funded by the Plants for Health and Wellness Initiative of the National Research Council Canada, The Pulse Research Network, funded by the Agricultural Bioproducts Innovation Program of Agriculture, and Agri-Food Canada, the Crop Genomics Initiative of Agriculture and Agri-Food Canada and the Ontario Research Fund, Research Excellence Program of the Ontario Ministry of Research and Innovation. Dengqun Liao was the recipient of a Visiting Fellowship in Canadian Government Laboratories of the Natural Sciences and Engineering Research Council of Canada. We thank



M. A. Afzal at the Pulses Research Centre, Bangladesh Agricultural Research Institute for the Barimash-2 seeds. We are grateful to staff at the Southern Crop Protection and Food Research Centre, Alex Molnar for preparation of figures, Ryan Austin for access to a Linux server and advice on bioinformatics software, Sudhakar Pandurangan for help with documenting plant

growth and bioinformatic analyses, Aga Pajak for technical assistance and Ralph Chapman and Mark Sumarah for preliminary S-methylhomoglutathione analyses. We acknowledge the help of Rey Interior at the Advanced Protein Technology Centre, Hospital for Sick Children, Toronto, ON, Canada, with amino acid analysis.

## REFERENCES

- Afzal, M. A., Bakr, M. A., Rahman, M. M., and Luna, N. K. (2002). Registration of “Barimash-2” blackgram. *Crop Sci.* 42, 985.
- Alvarez, C., Calo, L., Romero, L. C., Garcia, I., and Gotor, C. (2010). An O-acetylserine (thiol)lyase homolog with L-cysteine desulfhydrase activity regulates cysteine homeostasis in *Arabidopsis*. *Plant Physiol.* 152, 656–669.
- Amir, R., and Hacham, Y. (2008). “Methionine metabolism in plants,” in *Sulfur: A Missing Link Between Soils, Crops, and Nutrition*, ed. J. Jez (Madison: American Society of Agronomy, Crop Science Society of America, Soil Science Society of America), 251–279.
- Baldi, G., and Salamini, F. (1973). Variability of essential amino acid content in seeds of 22 *Phaseolus* species. *Theor. Appl. Genet.* 43, 75–78.
- Barberon, M., Berthomieu, P., Clairotte, M., Shibagaki, N., Davidian, J. C., and Gosti, F. (2008). Unequal functional redundancy between the two *Arabidopsis thaliana* high-affinity sulphate transporters SULTR1;1 and SULTR1;2. *New Phytol.* 180, 608–619.
- Barrera-Figueroa, B. E., Gao, L., Diop, N. N., Wu, Z., Ehlers, J. D., Roberts, P. A., et al. (2011). Identification and comparative analysis of drought-associated microRNAs in two cowpea genotypes. *BMC Plant Biol.* 11:127. doi:10.1186/1471-2229-11-127
- Baxter, I., Muthukumar, B., Park, H. C., Buchner, P., Lahner, B., Danku, J., et al. (2008). Variation in molybdenum content across broadly distributed populations of *Arabidopsis thaliana* is controlled by a mitochondrial molybdenum transporter (MOT1). *PLoS Genet.* 4:e1000004. doi:10.1371/journal.pgen.1000004
- Berardini, T. Z., Mundodi, S., Reiser, L., Huala, E., Garcia-Hernandez, M., Zhang, P., et al. (2004). Functional annotation of the *Arabidopsis* genome using controlled vocabularies. *Plant Physiol.* 135, 745–755.
- Bobb, A. J., Eiben, H. G., and Bustos, M. M. (1995). PvAlf, an embryo-specific acidic transcriptional activator enhances gene expression from phaseolin and phytohemagglutinin promoters. *Plant J.* 8, 331–343.
- Bourgis, F., Roje, S., Nuccio, M. L., Fisher, D. B., Tarczynski, M. C., Li, C. J., et al. (1999). S-Methylmethionine plays a major role in phloem sulfur transport and is synthesized by a novel type of methyltransferase. *Plant Cell* 11, 1485–1497.
- de Hoon, M. J., Imoto, S., Nolan, J., and Miyano, S. (2004). Open source clustering software. *Bioinformatics* 20, 1453–1454.
- Evans, I. M., and Boulter, D. (1975). S-Methyl-L-cysteine content of various legume meals. *Plant Foods Hum. Nutr.* 24, 257–261.
- Frendo, P., Jiménez, M. A. J. H., Mathieu, C., Duret, L., Gallesi, D., Van De Sype, G., et al. (2001). A *Medicago truncatula* homoglutathione synthetase is derived from glutathione synthetase by gene duplication. *Plant Physiol.* 126, 1706–1715.
- Galant, A., Preuss, M. L., Cameron, J. C., and Jez, J. M. (2011). Plant glutathione biosynthesis: diversity in biochemical regulation and reaction products. *Front. Plant Sci.* 2:45. doi:10.3389/fpls.2011.00045
- Giada, M. D. L. R., Miranda, M. T. M., and Marquez, U. M. L. (1998). Sulphur  $\gamma$ -glutamyl peptides in mature seeds of common beans (*Phaseolus vulgaris* L.). *Food Chem.* 61, 177–184.
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186.
- Goyer, A., Collakova, E., Shachar-Hill, Y., and Hanson, A. D. (2007). Functional characterization of a methionine gamma-lyase in *Arabidopsis* and its implication in an alternative to the reverse trans-sulfuration pathway. *Plant Cell Physiol.* 48, 232–242.
- Heeg, C., Kruse, C., Jost, R., Gutensohn, M., Ruppert, T., Wirtz, M., et al. (2008). Analysis of the *Arabidopsis* O-acetylserine (thiol)lyase gene family demonstrates compartment-specific differences in the regulation of cysteine synthesis. *Plant Cell* 20, 168–185.
- Hell, R., and Bergmann, L. (1990).  $\gamma$ -Glutamylcysteine synthetase in higher plants: catalytic properties and subcellular localization. *Planta* 180, 603–612.
- Hell, R., and Wirtz, M. (2011). Molecular biology, biochemistry and cellular physiology of cysteine metabolism in *Arabidopsis thaliana*. *Arabidopsis Book* 9, e0154.
- Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., et al. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 35, W585–587.
- Ikegami, F., Takayama, K., and Murakoshi, I. (1988a). Purification and properties of  $\beta$ -cyano-L-alanine synthase from *Lathyrus latifolius*. *Phytochemistry* 27, 3385–3389.
- Ikegami, F., Takayama, K., Tajima, C., and Murakoshi, I. (1988b). Purification and properties of  $\beta$ -cyano-L-alanine synthase from *Spinacia oleracea*. *Phytochemistry* 27, 2011–2016.
- Ikegami, F., Takayama, K., Kurihara, T., Horiuchi, S., Tajima, C., Shirai, R., et al. (1989). Purification and properties of  $\beta$ -cyano-L-alanine synthase from *Vicia angustifolia*. *Phytochemistry* 28, 2285–2291.
- Jander, G., and Joshi, V. (2010). Recent progress in deciphering the biosynthesis of aspartate-derived amino acids in plants. *Mol. Plant* 3, 54–65.
- Joshi, V., and Jander, G. (2009). *Arabidopsis* methionine  $\gamma$ -lyase is regulated according to isoleucine biosynthesis needs but plays a subordinate role to threonine deaminase. *Plant Physiol.* 151, 367–378.
- Kalavacharla, V., Liu, Z., Meyers, B. C., Thimmapuram, J., and Melmaie, K. (2011). Identification and analysis of common bean (*Phaseolus vulgaris* L.) transcriptomes by massively parallel pyrosequencing. *BMC Plant Biol.* 11:135. doi:10.1186/1471-2229-11-135
- Kasai, T., and Larsen, P. O. (1980). Chemistry and biochemistry of  $\gamma$ -glutamyl derivatives from plants including mushrooms (Basidiomycetes). *Fortschr. Chem. Org. Naturst.* 39, 173–285.
- Kasai, T., Shiroshita, Y., and Sakamura, S. (1986).  $\gamma$ -Glutamyl peptides of *Vigna radiata* seeds. *Phytochemistry* 25, 679–682.
- Kido, E. A., Barbosa, P. K., Neto, J. R., Pandolfi, V., Houllou-Kido, L. M., Crovella, S., et al. (2011). Identification of plant protein kinases in response to abiotic and biotic stresses using SuperSAGE. *Curr. Protein Pept. Sci.* 12, 643–656.
- Kim, W. S., Chronis, D., Juergens, M., Schroeder, A. C., Hyun, S. W., Jez, J. M., et al. (2012). Transgenic soybean plants overexpressing O-acetylserine sulfhydrylase accumulate enhanced levels of cysteine and Bowman-Birk protease inhibitor in seeds. *Planta* 235, 13–23.
- Klapheck, S. (1988). Homoglutathione: isolation, quantification and occurrence in legumes. *Physiol. Plant.* 74, 727–732.
- Klapheck, S., Zopes, H., Levels, H. G., and Bergmann, L. (1988). Properties and localization of the homoglutathione synthetase from *Phaseolus coccineus* leaves. *Physiol. Plant.* 74, 733–739.
- Kopriva, S., Mugford, S. G., Baraniecka, P., Lee, B. R., Matthewman, C. A., and Koprivova, A. (2012). Control of sulfur partitioning between primary and secondary metabolism in *Arabidopsis*. *Front. Plant Sci.* 3:163. doi:10.3389/fpls.2012.00163
- Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasisdharan, R., et al. (2012). The *Arabidopsis* Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* 40, D1202–D1210.
- Lee, M. S., Huang, T. F., Toro-Ramos, T., Fraga, M., Last, R. L., and Jander, G. (2008). Reduced activity of *Arabidopsis thaliana* HMT2, a methionine biosynthetic enzyme, increases seed methionine content. *Plant J.* 54, 310–320.
- Liao, D., Pajak, A., Karcz, S. R., Chapman, B. P., Sharpe, A. G., Austin, R. S., et al. (2012). Transcripts of sulphur metabolic genes are coordinately regulated in developing seeds of common bean lacking phaseolin and major lectins. *J. Exp. Bot.* 63, 6283–6295.

- Liu, F., Yoo, B.-C., Lee, J.-Y., Pan, W., and Harmon, A. C. (2006). Calcium-regulated phosphorylation of soybean serine acetyltransferase in response to oxidative stress. *J. Biol. Chem.* 281, 27405–27415.
- Liu, Z., Crampton, M., Todd, A., and Kalavacharla, V. (2012). Identification of expressed resistance gene-like sequences by data mining in 454-derived transcriptomic sequences of common bean (*Phaseolus vulgaris* L.). *BMC Plant Biol.* 12:42. doi:10.1186/1471-2229-12-42
- Marsolais, E., Pajak, A., Yin, F., Taylor, M., Gabriel, M., Merino, D. M., et al. (2010). Proteomic analysis of common bean seed with storage protein deficiency reveals up-regulation of sulfur-rich proteins and starch and raffinose metabolic enzymes, and down-regulation of the secretory pathway. *J. Proteomics* 73, 1587–1600.
- Moe, K. T., Chung, J. W., Cho, Y. I., Moon, J. K., Ku, J. H., Jung, J. K., et al. (2011). Sequence information on simple sequence repeats and single nucleotide polymorphisms through transcriptome analysis of mung bean. *J. Integr. Plant Biol.* 53, 63–73.
- Moran, J. F., Iturbe-Ormaetxe, I., Matamoros, M. A., Rubio, M. C., Clemente, M. R., Brewin, N. J., et al. (2000). Glutathione and homoglutathione synthetases of legume nodules. Cloning, expression, and subcellular localization. *Plant Physiol.* 124, 1381–1392.
- Mugford, S. G., Lee, B. R., Koprivova, A., Matthewman, C., and Kopriva, S. (2011). Control of sulfur partitioning between primary and secondary metabolism. *Plant J.* 65, 96–105.
- Mugford, S. G., Yoshimoto, N., Reichelt, M., Wirtz, M., Hill, L., Mugford, S. T., et al. (2009). Disruption of adenosine-5'-phosphosulfate kinase in *Arabidopsis* reduces levels of sulfated secondary metabolites. *Plant Cell* 21, 910–927.
- Nodari, R. O., Tsai, S. M., Gilbertson, R. L., and Gepts, P. (1993). Towards an integrated linkage map of common bean 2. Development of an RFLP-based linkage map. *Theor. Appl. Genet.* 85, 513–520.
- Noji, M., Inoue, K., Kimura, N., Gouda, A., and Saito, K. (1998). Isoform-dependent differences in feedback regulation and subcellular localization of serine acetyltransferase involved in cysteine biosynthesis from *Arabidopsis thaliana*. *J. Biol. Chem.* 273, 32739–32745.
- Otoul, E., Maréchal, R., Dardenne, G., and Desmedt, F. (1975). Des dipeptides souffrés différencient nettement *Vigna radiata* de *Vigna mungo*. *Phytochemistry* 14, 173–179.
- Padovese, R., Kina, S. M., Barros, R. M. C., Borelli, P., and Marquez, U. M. L. (2001). Biological importance of gamma-glutamyl-S-methylcysteine of kidney bean (*Phaseolus vulgaris* L.). *Food Chem.* 73, 291–297.
- Pandurangan, S., Pajak, A., Molnar, S. J., Cober, E. R., Dhaubhadel, S., Hernández-Sebastià, C., et al. (2012). Relationship between asparagine metabolism and protein concentration in soybean seed. *J. Exp. Bot.* 63, 3173–3184.
- Park, Y., Sheetlin, S., Ma, N., Madden, T. L., and Spouge, J. L. (2012). New finite-size correction for local alignment score distributions. *BMC Res. Notes* 5:286. doi:10.1186/1756-0500-5-286
- Peláez, P., Trejo, M. S., Iniguez, L. P., Estrada-Navarrete, G., Covarrubias, A. A., Reyes, J. L., et al. (2012). Identification and characterization of microRNAs in *Phaseolus vulgaris* by high-throughput sequencing. *BMC Genomics* 13:83. doi:10.1186/1471-2164-13-83
- Ranocha, P., Bourgis, F., Ziemak, M. J., Rhodes, D., Gage, D. A., and Hanson, A. D. (2000). Characterization and functional expression of cDNAs encoding methionine-sensitive and -insensitive homocysteine S-methyltransferases from *Arabidopsis*. *J. Biol. Chem.* 275, 15962–15968.
- Rathbun, W. B. (1967).  $\gamma$ -Glutamylcysteine synthetase from bovine lens. II. Cysteine analogue studies. *Arch. Biochem. Biophys.* 122, 73–84.
- Rébeillé, F., Jabrin, S., Bligny, R., Loizeau, K., Gambonnet, B., Van Wilder, V., et al. (2006). Methionine catabolism in *Arabidopsis* cells is initiated by a gamma-cleavage process and leads to S-methylcysteine and isoleucine syntheses. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15687–15692.
- Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178–183.
- Sehgal, C. B., and Gandhi, V. (1987). Histological and histochemical studies on the cotyledons of some legumes: I. changes during seed development. *Phytomorphology* 37, 9–16.
- Sekura, R., and Meister, A. (1977).  $\gamma$ -Glutamylcysteine synthetase. Further purification, “half of the sites” reactivity, subunits, and specificity. *J. Biol. Chem.* 252, 2599–2605.
- Shinmachi, F., Buchner, P., Stroud, J. L., Parmar, S., Zhao, F.-J., McGrath, S. P., et al. (2010). Influence of sulfur deficiency on the expression of specific sulfate transporters and the distribution of sulfur, selenium, and molybdenum in wheat. *Plant Physiol.* 153, 327–336.
- Smith, I. K., and Thompson, J. F. (1971). Purification and characterization of L-serine transacetylase and O-acetyl-L-serine sulphydrylase from kidney bean seedlings (*Phaseolus vulgaris*). *Biochim. Biophys. Acta* 227, 288–295.
- Tabé, L., Wirtz, M., Molvig, L., Droux, M., and Hell, R. (2010). Over-expression of serine acetyltransferase produced large increases in O-acetylserine and free cysteine in developing seeds of a grain legume. *J. Exp. Bot.* 61, 721–733.
- Takahashi, H., Kopriva, S., Giordano, M., Saito, K., and Hell, R. (2011). Sulfur assimilation in photosynthetic organisms: molecular functions and regulations of transporters and assimilatory enzymes. *Annu. Rev. Plant Biol.* 62, 157–184.
- Takahashi, H., Watanabe-Takahashi, A., Smith, F. W., Blake-Kalff, M., Hawkesford, M. J., and Saito, K. (2000). The roles of three functional sulphate transporters involved in uptake and translocation of sulphate in *Arabidopsis thaliana*. *Plant J.* 23, 171–182.
- Tan, Q., Zhang, L., Grant, J., Cooper, P., and Tegeder, M. (2010). Increased phloem transport of S-methylmethionine positively affects sulfur and nitrogen metabolism and seed development in pea plants. *Plant Physiol.* 154, 1886–1896.
- Taylor, M., Chapman, R., Beyaert, R., Hernández-Sebastià, C., and Marsolais, F. (2008). Seed storage protein deficiency improves sulfur amino acid content in common bean (*Phaseolus vulgaris* L.): redirection of sulfur from gamma-glutamyl-S-methyl-cysteine. *J. Agric. Food Chem.* 56, 5647–5654.
- Thompson, J. F., and Gering, R. K. (1966). Biosynthesis of S-methylcysteine in radish leaves. *Plant Physiol.* 41, 1301–1307.
- Tomatsu, H., Takano, J., Takahashi, H., Watanabe-Takahashi, A., Shibagaki, N., and Fujiwara, T. (2007). An *Arabidopsis thaliana* high-affinity molybdate transporter required for efficient uptake of molybdate from soil. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18807–18812.
- Vatamaniuk, O. K., Mari, S., Lu, Y. P., and Rea, P. A. (2000). Mechanism of heavy metal ion activation of phytochelatins (PC) synthase – blocked thiols are sufficient for PC synthase-catalyzed transpeptidation of glutathione and related thiol peptides. *J. Biol. Chem.* 275, 31451–31459.
- Walbot, V., Clutter, M., and Sussex, I. M. (1972). Reproductive development and embryogeny in *Phaseolus*. *Phytomorphology* 22, 59–68.
- Wang, C. S., and Vodkin, L. O. (1994). Extraction of RNA from tissues containing high levels of procyanidins that bind RNA. *Plant Mol. Biol. Rep.* 12, 132–145.
- Watanabe, M., Kusano, M., Oikawa, A., Fukushima, A., Noji, M., and Saito, K. (2008). Physiological roles of the  $\beta$ -substituted alanine synthase gene family in *Arabidopsis*. *Plant Physiol.* 146, 310–320.
- Watanabe, T., Tsugawa, M., Takayama, N., and Furukawa, Y. (1971). Changes of free amino acids of each organ in the development and growth of kidney bean plant. *J. Hokkaido Univ. Ed. Sect. II. A* 22, 45–54.
- Yi, H., and Jez, J. M. (2012). Assessing functional diversity in the soybean  $\beta$ -substituted alanine synthase enzyme family. *Phytochemistry* 83, 15–24.
- Yi, H., Juergens, M., and Jez, J. M. (2012). Structure of soybean  $\beta$ -cyanoalanine synthase and the molecular basis for cyanide detoxification in plants. *Plant Cell* 24, 2696–2706.
- Yi, H., Ravilious, G. E., Galant, A., Krishnan, H. B., and Jez, J. M. (2010). From sulfur to homoglutathione: thiol metabolism in soybean. *Amino Acids* 39, 963–978.
- Yin, F., Pajak, A., Chapman, R., Sharpe, A., Huang, S., and Marsolais, F. (2011). Analysis of common bean expressed sequence tags identifies sulfur metabolic pathways active in seed and sulfur-rich proteins highly expressed in the absence of phaseolin and major lectins. *BMC Genomics* 12:268. doi:10.1186/1471-2164-12-268
- Yoshimoto, N., Inoue, E., Watanabe-Takahashi, A., Saito, K., and Takahashi, H. (2007). Posttranscriptional regulation of high-affinity sulfate transporters in *Arabidopsis* by sulfur nutrition. *Plant Physiol.* 145, 378–388.
- Young, N. D., Debelle, F., Oldroyd, G. E. D., Geurts, R., Cannon, S. B., Udvardi, M. K., et al.

- (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480, 520–524.
- Zhai, J., Jeong, D. H., De Paoli, E., Park, S., Rosen, B. D., Li, Y., et al. (2011). MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes Dev.* 25, 2540–2553.
- Zheng, Y., Zhao, L., Gao, J., and Fei, Z. (2011). iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences. *BMC Bioinformatics* 12:453. doi:10.1186/1471-2105-12-453
- Zuber, H., Davidian, J. C., Aubert, G., Aimé, D., Belghazi, M., Lugan, R., et al. (2010). The seed composition of *Arabidopsis* mutants for the group 3 sulfate transporters indicates a role in sulfate translocation within developing seeds. *Plant Physiol.* 154, 913–926.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 08 January 2013; accepted: 04 March 2013; published online: 25 March 2013.
- Citation: Liao D, Cram D, Sharpe AG and Marsolais F (2013) Transcriptome profiling identifies candidate genes associated with the accumulation of distinct sulfur  $\gamma$ -glutamyl dipeptides in *Phaseolus vulgaris* and *Vigna mungo* seeds. *Front. Plant Sci.* 4:60. doi: 10.3389/fpls.2013.00060
- This article was submitted to *Frontiers in Plant Genetics and Genomics*, a specialty of *Frontiers in Plant Science*.
- Copyright © 2013 Liao, Cram, Sharpe and Marsolais. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

## APPENDIX

**Table A1 | Sulfate transporter and sulfur metabolic genes in the *P. vulgaris* genome.**

Gene name	Accession	Gene name	Accession
<i>Sultr1;1</i>	Phvul.006G207800	<i>MS4</i>	Phvul.006G007600
<i>Sultr1;2</i>	Phvul.009G028500	<i>MGL1</i>	Phvul.001G082000
<i>Sultr1;3</i>	Phvul.008G170800	<i>MGL2</i>	Phvul.004G090200
<i>Sultr2;1</i>	Phvul.009G028400	<i>GCL1</i>	Phvul.002G289200
<i>Sultr2;2</i>	Phvul.008G170700	<i>GCL2</i>	Phvul.002G157600
<i>Sultr2;3</i>	Phvul.001G250700	<i>hGS1</i>	Phvul.006G094500
<i>Sultr2;4</i>	Phvul.001G250800	<i>GS1</i>	Phvul.006G094600
<i>Sultr3;1</i>	Phvul.001G154200	<i>GS2</i>	Phvul.006G094700
<i>Sultr3;2</i>	Phvul.007G174100	<i>PCS1</i>	Phvul.001G162600
<i>Sultr3;3</i>	Phvul.004G161600	<i>PCS2</i>	Phvul.001G162700
<i>Sultr3;4</i>	Phvul.002G095200		
<i>Sultr3;5</i>	Phvul.002G095300		
<i>Sultr3;6</i>	Phvul.010G151000		
<i>Sultr3;7</i>	Phvul.005G171800		
<i>Sultr4;1</i>	Phvul.008G015600		
<i>Sultr5;1</i>	Phvul.010G152700		
<i>Sultr5;2</i>	Phvul.009G098800		
<i>Sultr5;3</i>	Phvul.001G056100		
<i>Sultr5;4</i>	Phvul.001G056300		
<i>APS1</i>	Phvul.007G062900		
<i>APS2</i>	Phvul.004G045400		
<i>APSK1</i>	Phvul.002G010900		
<i>APSK2</i>	Phvul.003G235300		
<i>APSK3</i>	Phvul.001G111000		
<i>APSR1</i>	Phvul.006G149200		
<i>APSR2</i>	Phvul.003G079800		
<i>SOX</i>	Phvul.004G054200		
<i>SIR</i>	Phvul.011G021800		
<i>SERAT1;1</i>	Phvul.001G170600		
<i>SERAT1;2</i>	Phvul.010G110600		
<i>SERAT2;1</i>	Phvul.006G055200		
<i>SERAT2;2</i>	Phvul.008G277800		
<i>SERAT3;2</i>	Phvul.002G114700		
<i>SERAT3;1</i>	Phvul.003G269000		
<i>BSAS1;1</i>	Phvul.002G045200		
<i>BSAS1;2</i>	Phvul.006G099100		
<i>BSAS1;3</i>	Phvul.007G057600		
<i>BSAS2;1</i>	Phvul.003G060200		
<i>BSAS3;1</i>	Phvul.008G061100		
<i>BSAS4;1</i>	Phvul.007G185200		
<i>BSAS4;2</i>	Phvul.007G185100		
<i>BSAS5;1</i>	Phvul.001G107500		
<i>CGS</i>	Phvul.008G038100		
<i>CBL</i>	Phvul.001G125400		
<i>HMT1</i>	Phvul.003G036100		
<i>HMT2</i>	Phvul.001G153900		
<i>HMT3</i>	Phvul.007G060300		
<i>MS1</i>	Phvul.001G113800		
<i>MS2</i>	Phvul.004G051600		
<i>MS3</i>	Phvul.005G048200		

(Continued)