

Systematic Comparisons of Orthologous Selenocysteine Methyltransferase and Homocysteine Methyltransferase Genes from Seven Monocots Species

De-yong ZHAO^{1,2,4*}, Fu-lai SUN³, Bo ZHANG¹, Zhi-qiang ZHANG², Long-quan YIN²

¹Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Plateau Institute of Biology, Chinese Academy of Sciences, No. 23, Xinning Road, Xining 810001, Qinghai, P. R. China; zdy1984@126.com (*corresponding author)

²Binzhou Key Laboratory of Wheat Breeding, No.68, 2nd Fenghuang Road, Bincheng District, Binzhou 256651, Shandong, P.R.China

³Binzhou Seed Station, No. 12, Huanghe Road, Binzhou, Shandong, P.R.China

⁴University of Chinese Academy of Sciences, No. 19 (Jia), Shijingshan District, Beijing 100049, P. R. China

Abstract

Identifying and manipulating genes underlying selenium metabolism could be helpful for increasing selenium content in crop grain, which is an important way to overcome diseases resulted from selenium deficiency. A reciprocal smallest distance algorithm (RSD) approach was applied using two experimentally confirmed Homocysteine S-Methyltransferases genes (HMT1 and HMT2) and a putative Selenocysteine Methyltransferase (SMT) from dicots plant *Arabidopsis thaliana*, to explore their orthologs in seven sequenced diploid monocot species: *Oryza sativa*, *Zea mays*, *Sorghum bicolor*, *Brachypodium distachyon*, *Hordeum vulgare*, *Aegilops tauschii* (the D-genome donor of common wheat) and *Triticum urartu* (the A-genome donor of common wheat). HMT1 was apparently diverged from HMT2 and most of SMT orthologs were the same with that of HMT2 in this study, leading to the hypothesis that SMT and HMT originate from one common ancestor gene. Identifying orthologs provide candidates for further experimental confirmation; also it could be helpful in designing primers to clone SMT or HMT orthologs in other crops.

Keywords: Homocysteine S-Methyltransferases, ortholog, selenium, Selenocysteine Methyltransferase

Introduction

Selenium is an important micronutrient, essential for both humans and animals (Schwarz and Foltz, 1957; Schrauzer and Surai, 2009) as certain proteins require selenocysteine in their active site (Stadtman, 1990; 1996). Selenium deficiency can lead to cancer, heart disease, hypothyroidism, a weakened immune system and Kaschin-Beck disease (Chen *et al.*, 1980; Mo, 1987; Peng and Yang, 1991). Therefore, it is of great importance to increase selenium content in crops. Scientists have tried to increase grain selenium content through fertiliser-application fortification or breeding selenium-rich crop using the genotypic variation (Hawkesford *et al.*, 2007; Lyons *et al.*, 2005). However, there was little genotypic variation and much of the effects were associated with selenium spatial variation in soil (Lyons *et al.*, 2005). Consequently, manipulating expression of genes underlying selenium metabolism through genetic modification could be a useful approach for increasing selenium content in crop grain. The prerequisite of genetic modification approach requires identifying the genes involved in selenium metabolism and clarifying their contributing roles to the final selenium content in plants.

The final selenium content in different plant organs is controlled by three processes, e.g. uptake of selenium from soil, assimilation of selenium and translocation into different organs.

Therefore, understanding the molecular mechanisms of selenium uptake, assimilation and translocation, will facilitate further the understanding of how could an increase of selenium content in grains of cereal crops be achieved. It is generally accepted that plant absorb selenate by sulphate transporters due to the similarity between sulphate and selenate (Broadley *et al.*, 2006; Ellis and Salt, 2003; Hawkesford, 2003; Li *et al.*, 2008). By contrast, the uptake mechanism of another major selenium form in soil, selenite, remains to be illustrated (Li *et al.*, 2008; Terry *et al.*, 2000). After being absorbed into plant, selenate and selenite are subjected to a series of enzymatic catalyse (Ellis and Salt, 2003; Rayman, 2008; Sors *et al.*, 2005; Tagmount *et al.*, 2002; Whanger, 2004). Selenocysteine Methyltransferase (SMT) has been indicated as playing a crucial role in selenium accumulation in selenium accumulator *Astragalus bisulcatus*, which often grows in a high selenium habitat (Neuhierl *et al.*, 1999; Sors *et al.*, 2009). Overexpression of SMT can increase selenium accumulation in *Arabidopsis* and Indian mustard (LeDuc *et al.*, 2004). SMT shares high identity with Homocysteine S-Methyltransferase (HMT). Until now, there is little experimental knowledge about SMT function in the major cereal crops like rice, maize and common wheat. although their genomes have been sequenced (Brenchley *et al.*, 2012; Goff *et al.*, 2002; Schnable *et al.*, 2009; Yu *et al.*, 2002). As selenium is

tightly associated with human health and many studies support its protective role against various types of cancer (Naithani, 2008), overexpression of exogenous SMT with high capability in crops such as wheat, corn and rice to enhance the Methyl-Seleno Cysteine (MeSeCys) content may be helpful to overcome selenium deficiency diseases.

With expansion of available genome sequence data, it is feasible to explore orthologs in crop genomes using an experimentally confirmed SMT or HMT from model plant through Blast approach. Reciprocal best hit (RBH) approach is a widely used method to identify putative orthologs across two genomes (Bork et al., 1998; Plett et al., 2010; Tatusov et al., 1997). However, this Blast search often returns as the highest scoring hit that is not the nearest phylogenetic neighbour of the query sequence (Koski and Golding, 2001). Meanwhile, reciprocal best hit may wrongly exclude an authentically orthologous pair from consideration (Wall et al., 2003). Reciprocal smallest distance algorithm (RSD) which is improved upon RBH can find putative orthologs missed by RBH as it is less likely to be misled by the presence of a close paralog (Wall et al., 2003). Therefore, RSD approach was applied in this study using two experimentally confirmed HMT genes (AtHMT1 and AtHMT2) and a putative SMT (AtSMT) from dicot plants *Arabidopsis thaliana* to retrieve their orthologs in seven sequenced diploid monocot species including rice (*Oryza sativa*), maize (*Zea mays*), *Sorghum bicolor*, *Brachypodium distachyon*, barley (*Hordeum vulgare*), *Triticum urartu* and *Aegilops tauschii*. Among these seven monocots, rice and maize are two major crops worldwide: *T. urartu* and *A. tauschii* belonging to the tribe Triticeae in Poaceae and are the progenitor of allohexaploid common wheat A and D genome respectively (Jia et al., 2013; Ling et al., 2013); barley is one of experimental models for Triticeae biology (Schulte et al., 2009), while *Sorghum bicolor* and *Brachypodium distachyon* are also useful references to study evolutionary relationships among monocots, as these two genomes has been sequenced (Paterson et al., 2009; The International Brachypodium Initiative, 2010).

Evolution of single or several genes, e.g. mutation in nucleotide sequence, would change one or some traits of a species, hence contribute to phenotypic evolution. As an example, the Selenocysteine Methyltransferase from non-accumulator *Astragalus drummondii* shares high homology with SMT from accumulator *Astragalus bisulcatus*, but lacks Selenocysteine Methyltransferase activity *in vitro*, explaining why there is little or no detectable levels of MeSeCys in the non-accumulator plant compared to MeSeCys as the predominant form in its relative species *Astragalus bisulcatus* (Sors et al., 2009). High similarity between Homocysteine Methyltransferase and Cysteine Methyltransferase, but different enzyme specificity draws our interest to study evolutionary relationship between these two enzymes among the seven monocots.

Materials and methods

Two experimentally confirmed Selenocysteine Methyltransferase (SMT) genes in *Astragalus bisulcatus*, SMT (Genbank accession number, CAA10368) (Neuhierl et al., 1999) and BoSMT (Genbank accession number, AAX20123.1.) (Lyi et al., 2005), were used to retrieve the orthologs from *A. thaliana* through Blastp approach. Since

both of them showed the shortest pairwise distance with a putative SMT gene (GenBank accession number, BAC42654.1) in *A. thaliana*, BAC42654.1 was further used as query sequence in RSD. Protein sequences of two experimentally confirmed Homocysteine S-Methyltransferases, AtHMT1 (accession number AAF23821.1) and AtHMT2 (AAF23822.1) (Ranocha et al., 2000) from *A. thaliana*, were used as query sequences for exploring HMT1 and HMT2 orthologs.

A Blastp programme using query sequences was performed to search orthologs of maize, rice, barley, *S. bicolor*, *B. distachyon*, *T. urartu* and *A. tauschii* in the public database (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The set of hits exceeding a significance threshold $E < 10^{-20}$ and the query coverage $\geq 80\%$ was obtained for each round forward Blastp. The sequences obtained were used as queries in subsequent reverse Blastp searches against the *Arabidopsis* database; sequences which could return the original sequence in the reverse Blastp were used for further analysis. Subsequently, the alignment of these sequences obtained and the pairwise distance were calculated using MEGA5 (Tamura et al., 2011); sequences showing the shortest distance with query sequence were then used in the second round Blastp, so that a set of sequences was obtained. Pairwise distance between these sequences in the second round Blastp and the sequence showing shortest distance with query sequence were calculated, and the pairs which showed the smallest distance were designated as orthologs. Comparison between rice, barley, maize, *S. bicolor*, *B. distachyon*, *T. urartu* and *A. tauschii* were conducted by the same method mentioned above, where query sequences were those which had been designated as HMT or SMT orthologs in the reciprocal Blastp between *A. thaliana* and each of these seven monocots.

Evolutionary analysis were conducted in MEGA5 (Tamura et al., 2011) using the Minimum Evolution method (Rzhetsky and Nei, 1992). Evolutionary tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the poisson correction method (Zuckerkanndl and Pauling, 1965) and were in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated. The minimum evolution tree was searched using the Close-Neighbour-Interchange (CNI) algorithm (Nei and Kumar, 2000) at a search level of 1. The Neighbour-joining algorithm (Saitou and Nei, 1987) was used to generate the initial tree.

Results

Putative SMT orthologs

Seven putative SMT orthologs from rice, maize, barley, *S. bicolor* and *B. distachyon*, *T. urartu* and *A. tauschii* were identified using putative *A. thaliana* SMT (BAC42654.1) as initial query sequence (Fig. 1). Reverse Blast towards *A. thaliana* genome, however, returned three different orthologs, NP 191884.1, AAF23822.1 and AAM 65096.1, rather than BAC42654.1. Similar situations also occurred when RSD were conducted among these seven monocots species. Besides the putative SMT (BAC42654.1), three other paralogs were also found in *A. thaliana* genome. Five paralogs were found in maize, three in rice,

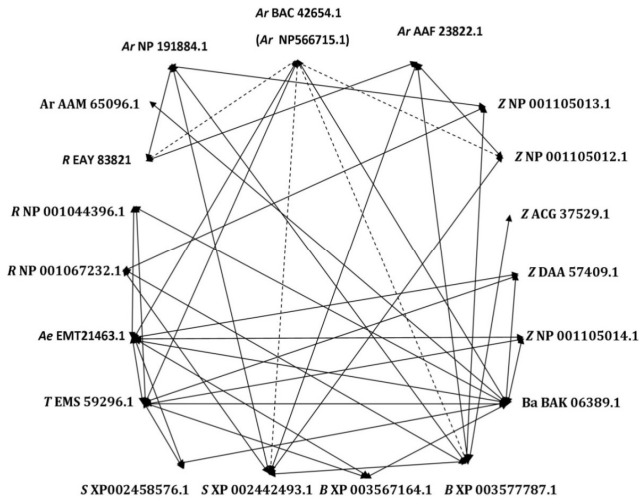


Fig. 1. Reciprocal Blast among *A. thaliana* (Ar), rice (R), *Z. mays* (Z), *S. bicolor* (S), barley (Ba), *B. distachyon* (B), *A. tauschii* (Ae) and *T. urartu* (T) genome utilizing AtSMT as query sequence

The symbol with bold line and two arrows ' \leftrightarrow ' indicates the shortest distance through reciprocal shortest distance Blast approach, while open arrows with dashed line ' \dashrightarrow ' indicates seven orthologs in rice, maize, *S. bicolor*, *B. distachyon*, barley, *A. tauschii* and *T. urartu* that were identified in the first round forward Blast.

two in *S. bicolor* and *B. distachyon*. By contrast, only one ortholog was found for barley, *A. tauschii* and *T. urartu* respectively. This implied gene duplication event of the putative SMT which might have occurred in *A. thaliana*, maize, rice, *S. bicolor* and *B. distachyon*.

Phylogenetic tree was further constructed and it can be divided into three groups (Fig. 2). Within group one, there were three orthologs from maize, two from rice, one from *S. bicolor* and one from *B. distachyon*. Orthologs from *A. thaliana* constituted as group two. Within group three, there were two orthologs from maize, one from each of *S. bicolor*, rice, *B. distachyon*, barley, *T. urartu* and *A. tauschii*. There was a very close genetic relationship between maize orthologs NP 001105013.1 and ACG 37579.1; rice orthologs EAY 83821 and NP 001067232.1; *A. thaliana* orthologs NP 191884.1 and AAF 23822.1. Further alignment demonstrated ACG 37579.1 (383 amino acids) was with 17 amino acids longer than NP 001105013.1 (366 amino acids), 365 amino acids of NP 001105013.1 was the same with that of ACG 37579.1 except the last amino acid "A" substituted by "G" in ACG 37579.1. Two rice orthologs EAY 83821 and NP 001067232.1 were the same except that the amino acid "M" in position 269 of EAY 83821 is substituted by "I" in NP 001067232.1, while *A. thaliana* orthologs NP 191884.1 and AAF 23822.1 were the same except that the amino acid "T" in position 103 of NP 191884.1 was substituted by "C" in AAF 23822.1.

HMT1 orthologs

One ortholog from each genome of rice, barley, maize, *S. bicolor* and *B. distachyon*, *A. tauschii* and *T. urartu*, were found by first round Blast using AtHMT1 (Genbank accession number, AAF23821.1), reverse Blast returned a sequence with the accession number NP 189219.1 (Fig. 3). Sequence alignment analysis showed these two sequences, AAF23821.1 and NP 189219.1 were identical. After two rounds RSD Blast,

the reciprocal Blast diagram among *A. thaliana*, rice, barley, maize, *S. bicolor*, *B. distachyon*, *A. tauschii* and *T. urartu* genomes was constructed. Different from the result of SMT, no HMT1 paralog was found in maize, rice, *S. bicolor* and *B. distachyon*.

The evolutionary tree constructed with HMT1 orthologs can be classified into two apparent groups: AtHMT1 (AAF23821.1/ NP 189219.1) constituted as one group, while orthologs from the other seven genomes another group (Fig. 4). This implied the large difference between dicots *A. thaliana* and the seven monocot species. Orthologs from maize (NP 001105011.1) and *S. bicolor* (XP 002468259.1) displayed a

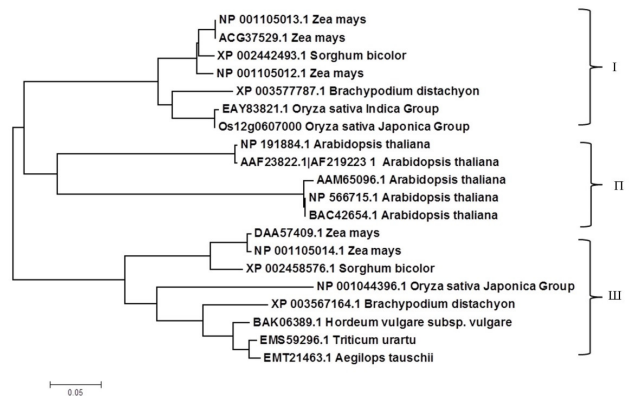


Fig. 2. Evolutionary relationships of SMT orthologs

The analysis was involved in 291 positions for 20 amino acid sequences. The optimal tree had the sum of branch length = 1.45.

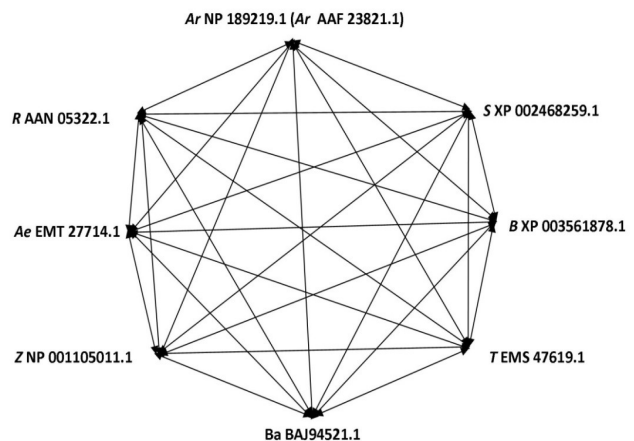


Fig. 3. Reciprocal Blast among *A. thaliana* (Ar), rice (R), *Z. mays* (Z), *S. bicolor* (S), barley (Ba), *B. distachyon* (B), *A. tauschii* (Ae) and *T. urartu* (T) genomes utilizing AtHMT1 as query sequence

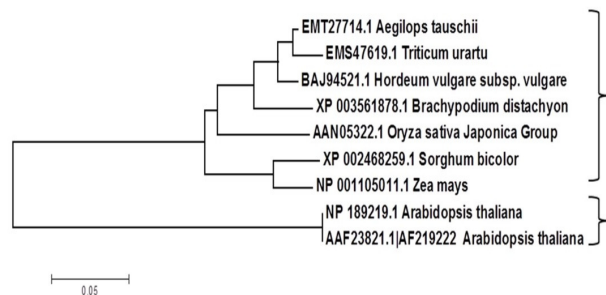


Fig. 4. Evolutionary relationships of AtHMT1 orthologs

The analysis was involved in 319 positions for nine amino acid sequences. The optimal tree had the sum of branch length = 0.61

closer evolutionary relationship, while orthologs from Triticeae species *A. tauschii*, *T. urartu* and barley clustered closer.

HMT2 orthologs

First round RSD Blast using AtHMT2 as query sequence returned three maize paralogs (Fig. 5), while only one was found in rice, barley, *S. bicolor*, *B. distachyon*, *A. tauschii* and *T. urartu*. Second round Blast using these three maize paralogs towards *A. thaliana* genome returned another *A. Thaliana* paralog (NP 001064628.1). Similarly, reciprocal Blast between rice and three other genomes of maize, *S. bicolor*, *B. distachyon*, using initial sequence NP_001064628, returned a different rice paralog NP_001067232.1. Reciprocal Blastp result between *A. tauschii* and rice, maize, *B. distachyon*, *S. bicolor* genome using EMT 21463.1 was the same as that using *T. urartu* ortholog EMS 59296.1. These two orthologs also clustered tightly in the evolutionary tree (Fig. 6), implying a closer evolutionary relationship.

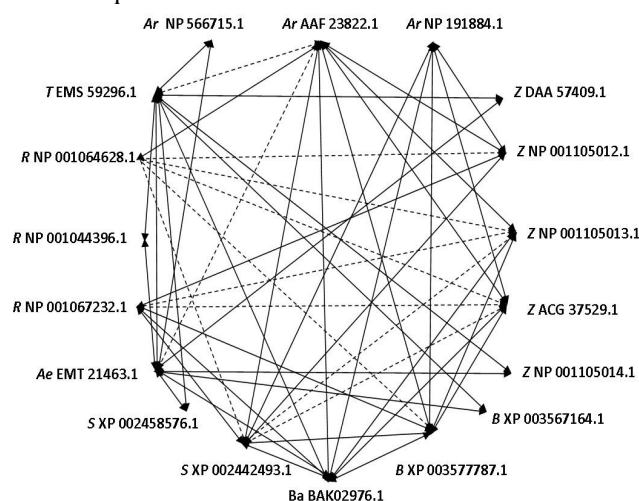


Fig. 5. Reciprocal Blast among *A. thaliana* (Ar), rice (R), *Z. mays* (Z), *S. bicolor* (S), barley (Ba), *B. distachyon* (B), *A. tauschii* (Ae) and *T. urartu* (T) genomes using AtHMT2 as query sequence

The orthologs can be subdivided into four apparent groups (Fig. 6). Within group one, three maize orthologs (NP 001105012.1, NP 001105013.1, ACG 37529.1) clustered tightly with one *S. bicolor* ortholog (XP 002442493.1), while one ortholog from *B. distachyon* (XP 003577787.1) clustered tightly with one from barley (BAK 02976.1). Rice ortholog NP 001064628.1 constituted as group two, while orthologs from *A. thaliana* constituted as group three. Within group four, there were the other two maize orthologs (DAA 57409.1, NP 001105014.1) along with one *S. bicolor* ortholog (XP 002458576.1) and one *Brachypodium* ortholog (XP 003567164.1).

Evolutionary relationship and conserved amino acids between HMT and SMT

High identity between SMT and HMT confers common function catalysing cysteine into methylcysteine, but different affinity capability (Sors et al., 2009); Orthologs of HMT1 were apparently diverged from that of HMT2 and SMT (Figs. 7 and 8). These two facts lead to the hypothesis that

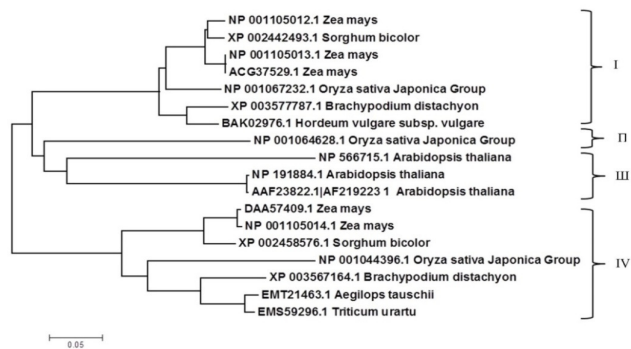


Fig. 6. Evolutionary relationships of AtHMT2 orthologs. The analysis was involved in 287 positions for 18 amino acid sequences. The optimal tree had the sum of branch length = 1.63.

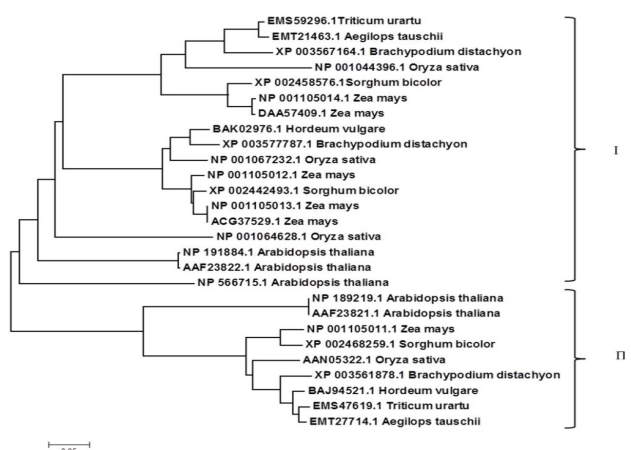


Fig. 7. Evolutionary tree of HMT1 and HMT2 orthologs. The optimal tree had the sum of branch length = 2.38 as shown hereby. The tree was drawn to scale. The analysis involved 27 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 284 positions in the final dataset.

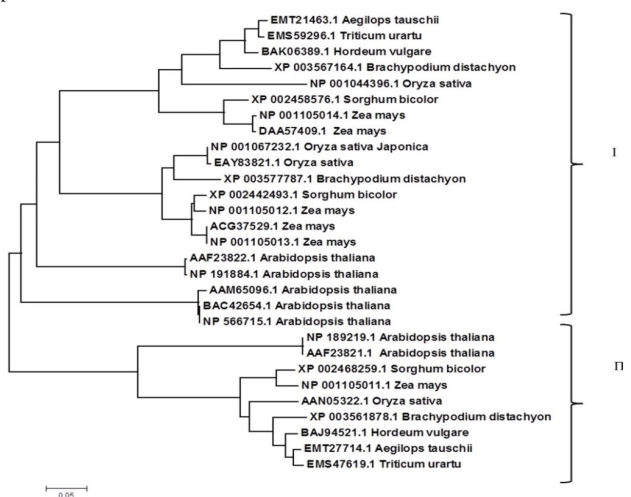


Fig. 8. Evolutionary tree of HMT1 and SMT orthologs. The optimal tree had the sum of branch length = 2.24 as shown hereby. The tree was drawn to scale. The analysis involved 29 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 288 positions in the final dataset.

SMT and HMT may originate from one common ancestor. It was further supported by that most (16 out of 20) orthologs are the same between SMT (Fig. 1) and HMT2 (Fig. 2) in this study. Different conserved amino acids which might be responsible for the different enzyme catalysing capability between HMT and SMT were found (Table 1) by

Table 1. Different conserved amino acids in HMT1 and SMT

Position	Amino acid in SMT	Amino acid in HMT1
25	R/T/K	E
60	V/A	K/R/S
87	A	S
94	Q/E	P
125	A/L/V/Y	W
147	R/Q/K/H	Y
148	R/H/P	N
235	K	V
238	V/A	E/K
270	R/G	Q
277	L/I	R/C
298	Y	W
326	R/N/T/K	A
334	G/A	C
335	A/V/G/E	F/L
348	V/A	E/D
349	S	M/V/A
351	V	A

sequence alignment involving 20 SMT orthologs and 9 HMT1 orthologs. For example, at position 87, amino acid 'A' in SMT was replaced by 'S' in HMT1. At some positions in both enzymes, the amino acid is not highly conserved, but shows preference. For example, at position 60, the preference amino acid for SMT is 'V' or 'A', while for HMT1 is 'K' or 'R' or 'S'.

Discussions

Adaptive mutation might have occurred in selenium accumulators as a strategy to adapt to high selenium environment. Selenocysteine Methyltransferases with high affinity towards L-Selenocysteine may be only one of these adaptive mechanisms. Besides Selenocysteine Methyltransferases, some other candidate genes were also proposed accounting for the three QTLs (Quantitative Trait Loci) on chromosomal 1, 3, 5 for selenium tolerance in *Arabidopsis thaliana* (Zhang et al., 2006). Further comparison on mechanisms of selenium metabolism between Selenium non-accumulators and accumulators may lead to discovery of new regulatory mechanism sustaining selenium accumulators to accumulate selenium.

Quantitative Trait Loci (QTL) mapping is a powerful and traditional approach to identify QTLs that contribute to phenotypic variation; for example, three QTL loci on chromosomal 1, 3, 5 in *Arabidopsis thaliana* (Zhang et al., 2006) and six QTL loci in rice (Norton et al., 2010) were found accounting for grain selenium content using this approach. However, exploring novel genes/factors underlying selenium metabolism in crops by traditional QTL mapping seems to be difficult for the mapping population grown in the field, due to heterogeneous distribution of selenium in soil and the largely dependence of grain selenium content upon soil selenium content. Therefore, towards the goal of understanding molecular mechanism of selenium metabolism and producing genetic modified crops with enhanced selenium accumulating ability, manipulating an experimentally confirmed selenium metabolic gene from other plants in crops may serve as an alternative way.

MeSeCys accounts as the predominant form in selenium accumulators while the content is lower in non-selenium accumulator, suggesting that SMT plays an important role in

selenium accumulation. The predominant form of selenium in non-selenium accumulator cereal is Selenomethionine (Whanger, 2004); for example, Selenomethionine in wheat grain accounts 56-83% of the total selenium content, followed by SeCys and MeSeCys accounting 4-12% and 1-4%, respectively (Whanger, 2002). Hence it is reasonable to deduce that the activity of Selenocysteine Methyltransferase in wheat might be lower than that in accumulators. Moreover, based on the fact the catalysing activity of *Astragalus bisulcatus* Selenocysteine Methyltransferase towards L-Selenocysteine was much higher than that of L-Cysteine (Neuhierl and Böck 1996), it can be hypothesised that Selenocysteine Methyltransferase in non-selenium accumulators such as wheat, has lower capability to catalyse L-Selenocysteine.

Although Selenocysteine Methyltransferase and Homocysteine Methyltransferase have high sequence similarity, they do have apparent different roles in selenium metabolism. Previous study of Sors et al. (2009) indicated that a putative SMT enzyme from the non-accumulator *Astragalus drummondii* showed a high degree of homology with the accumulator *Astragalus bisulcatus* SMT (AbSMT), but lacked the Selenocysteine Methyltransferase activity *in vitro*; also Ala to Thr amino acid mutation at the predicted active site of AbSMT resulted in a new enzymatic capacity to methylate homocysteine and exhibit a six fold higher capacity to methylate selenocysteine, indicating that SMT mutation can affect its enzyme activity.

Specific amino acids belonging to SMT or HMT1 orthologs were also found in this study. These facts support the hypothesis that SMT originate from mutation during duplication process of HMT. It is also necessary to experimentally confirm the substrate specificity of novel Selenocysteine Methyltransferases or Homocysteine Methyltransferases which are cloned using degenerate primers based on conserved sequences.

As selenium is tightly associated with human health and many studies support its protective role against various types of cancer (Naithani, 2008), overexpression of exogenous SMT with high capability to methyl Selenocysteine in crops such as wheat, corn and rice, in order to enhance the MeSeCysteine content, may be helpful in targeting selenium deficiency diseases and preventing some kind of cancers.

Conclusions

Orthologs for both Selenocysteine Methyltransferases and Homocysteine Methyltransferases (HMT1, HMT2) of seven monocot species were obtained by the effective RSD method in this study. Systematic comparison among these orthologs and bioinformatics analysis were subsequently conducted. SMT and HMT might originate from one common ancestor gene since HMT1 was apparently diverged from HMT2 and most of SMT orthologs were the same with that of HMT2 in this study. It is also necessary to experimentally confirm the substrate specificity of novel Selenocysteine Methyltransferases or Homocysteine Methyltransferases, especially those cloned by degenerate primers based on conserved sequences, since these two genes shares high identity. Identified orthologs in this study provide candidates for further experimental confirmation; also could be helpful in designing primers to clone SMT or HMT orthologs in other crops.

Acknowledgements

The research was financially supported by National Natural Science Foundation of China (31101140), the Applied Basic Research Project in Qinghai Province (2011-Z-716) and the Program of West Action of CAS (KZCX2-XB3-05).

References

- Bork P, Dandekar T, Diaz-Lazcoz Y, Eisenhaber F, Huynen M, Yuan Y (1998). Predicting function: from genes to genomes and back. *J Mol Biol* 283(4):707-725.
- Brenchley R, Spannagl M, Pfeifer M, Barker GLA, D'Amore R, Allen A, McKenzie N, Kramer M, et al. (2012). Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* 491(7426):705-710.
- Broadley MR, Whitea PJ, Bryson RJ, Meacham MC, Bowen HC, Johnson SE, Hawkesford MJ, McGrath SP, Zhao FJ, Breward N, Harriman M, Tuckera M (2006). Biofortification of UK food crops with selenium. *Proc Nutr Soc* 65(2):169-181.
- Chen X, Yang G, Chen J, Chen X, Wen Z, Ge K (1980). Studies on the relations of selenium and Keshan disease. *Biol Trace Elem Res* 2(2):91-107.
- Clark LC, Combs GF Jr, Turnbull BW, Slate EH, Chalker DK, Chow J, Davis LS, Glover RA, et al. (1996). Effects of selenium supplementation for cancer prevention in patients with carcinoma of the skin: A Randomized Controlled Trial. *JAMA* 276(24):1957-1963.
- Combs GF Jr (2000). Food system-based approaches to improving micronutrient nutrition: The case for selenium. *BioFactors* 12:39-43.
- Ellis DR, Salt DE (2003). Plants, selenium and human health. *Curr Opin Plant Biol* 6(3):273-279.
- Goff SA, Ricke D, Lan T, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, et al. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). *Science* 296(5565):92-100.
- Hawkesford MJ (2003). Transporter gene families in plants: the sulphate transporter gene family-redundancy or specialization? *Physiol Plantarum* 117(2):155-163.
- Hawkesford MJ, Zhao FJ (2007). Strategies for increasing the selenium content of wheat. *J Cereal Sci* 46(3):282-292.
- Jia J, Zhao S, Kong X, Li Y, Zhao G, He W, Appels R, Pfeifer M, et al. (2013). *Aegilops tauschii* draft genome sequence reveals agene repertoire for wheat adaptation. *Nature* 496:91-95.
- Koski LB, Golding GB (2001). The closest BLAST hit is often not the nearest neighbor. *J Mol Evol* 52(6):540-542.
- LeDuc DL, Tarun AS, Montes-Bayon M, Mejia J, Malit MF, Wu CP, AbdelSamie M, Chiang CY, et al. (2004). Overexpression of selenocysteine methyltransferase in Arabidopsis and Indian mustard increases selenium tolerance and accumulation. *Plant Physiol* 135(1):377-383.
- Li HF, McGrath SP, Zhao FJ (2008). Selenium uptake, translocation and speciation in wheat supplied with selenate or selenite. *New Phytol* 178(1):92-102.
- Ling HQ, Zhao S, Liu D, Wang J, Sun H, Zhang C, Fan H, Li D, et al. (2013). Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature* 496:87-90.
- Lyi SM, Heller LI, Rutzke M, Welch RM, Kochian LV, Li L (2005). Molecular and biochemical characterization of the selenocysteine Se-methyltransferase gene and Se-methylselenocysteine synthesis in broccoli. *Plant Physiol* 138(1):409-420.
- Lyons G, Ortiz-Monasterio I, Stangoulis J, Graham R (2005). Selenium concentration in wheat grain: Is there sufficient genotypic variation to use in breeding? *Plant Soil* 269(1-2):369-380.
- Mo D (1987). Pathology and selenium deficiency in Kaschin-Beck disease. Source: Selenium in biology and medicine: proceedings of the Third International Symposium on Selenium in Biology and Medicine, held May 27-June 1, 1984, pp 924-933, ISBN 04-422-21088.
- Naithani R (2008). Organoselenium compounds in cancer chemoprevention. *Mini-Rev Med Chem* 8(7):657-668.
- Nei M, Kumar S (2000). Molecular evolution and phylogenetics, Oxford University Press, USA.
- Neuhierl B, Böck A (1996). On the mechanism of selenium tolerance in selenium - accumulating plants. *Eur J Biochem* 239(1):235-238.
- Neuhierl B, Thanbichler M, Lottspeich F, Böck A (1999). A family of S-Methylmethionine-dependent Thiol/Selenol Methyltransferases role in selenium tolerance and evolutionary relation. *J Biol Chem* 274(9):5407-5414.
- Norton GJ, Deacon CM, Xiong L, Huang S, Meharg AA, Price AH (2010). Genetic mapping of the rice ionome in leaves and grain: identification of QTLs for 17 elements including arsenic, cadmium, iron and selenium. *Plant Soil* 329(1-2):139-153.
- Paterson A, Bowers J, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, et al. (2009). The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457(7229):551-556.
- Peng A, Yang C (1991). Examination of the roles of selenium in the Kaschin-Beck disease. *Biol Trace Elem Res* 28(1):1-9.
- Plett D, Toubia J, Garnett T, Tester M, Kaiser BN, Baumann U (2010). Dichotomy in the NRT gene families of dicots and grass species. *PLoS one* 5(12):e15289.
- Ranocha P, Bourgis F, Ziemak MJ, Rhodes D, Gage DA, Hanson AD (2000). Characterization and functional expression of cDNAs encoding methionine-sensitive and-insensitive homocysteine S-methyltransferases from Arabidopsis. *J Biol Chem* 275(21):15962-15968.
- Rayman MP (2008). Food-chain selenium and human health: emphasis on intake. *Brit J Nutr* 100(2):254-268.
- Rzhetsky A, Nei M (1992). A simple method for estimating and testing minimum-evolution trees. *Mol Biol Evol* 9(5):945-967.
- Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4(4):406-425.
- Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326(5956):1112-1115.
- Schrauzer GN, Surai PF (2009). Selenium in human and animal nutrition: Resolved and unresolved issues. A partly historical treatise

- in commemoration of the fiftieth anniversary of the discovery of the biological essentiality of selenium, dedicated to the memory of Klaus Schwarz (1914-1978) on the occasion of the thirtieth anniversary of his death. *Crit Rev Biotechnol* 29(1):2-9.
- Schulte DTJ, Close, Graner A, Langridge P, Matsumoto T, Muehlbauer G, Sato K, Schulman A, Waugh R, Wise RP, Stein N (2009). The international barley sequencing consortium-at the threshold of efficient access to the barley genome. *Plant Physiol* 149(1):142-147.
- Schwarz K, Foltz CM (1957). Selenium as an integral part of factor 3 against dietary necrotic liver degeneration. *J Am Chem Soc* 79(12):3292-3293.
- Sors TG, Ellis DR, Salt DE (2005). Selenium uptake, translocation, assimilation and metabolic fate in plants. *Photosynth Res* 86(3):373-389.
- Sors TG, Martin CP, Salt DE (2009). Characterization of selenocysteine methyltransferases from *Astragalus* species with contrasting selenium accumulation capacity. *Plant J* 59(1):110-122.
- Stadtman TC (1990). Selenium biochemistry. *Annu Rev Biochem* 59:111-127.
- Stadtman TC (1996). Selenocysteine. *Annu Rev Biochem* 65:83-100.
- Tagmount A, Berken A, Terry N (2002). An essential role of S-Adenosyl-L-Methionine: L-MethionineS-Methyltransferase in selenium volatilization by plants. Methylation of Selenomethionine to Selenium-Methyl-L-Selenium-Methionine, the precursor of volatile selenium. *Plant Physiol* 130(2):847-856.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10):2731-2739.
- Tatusov RL, Koonin EV, Lipman DJ (1997). A genomic perspective on protein families. *Science* 278(5338):631-637.
- Terry N, Zayed A, de Souza MP, Tarun AS (2000). Selenium in higher plants. *Annu Rev Plant Biol* 51(1):401-432.
- The International Brachypodium Initiative (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463(7282):763-768.
- Wall DP, Fraser HB, Hirsh AE (2003). Detecting putative orthologs. *Bioinformatics* 19(13):1710-1711.
- Whanger P (2002). Selenocompounds in plants and animals and their biological significance. *J Am Coll Nutr* 21(3):223-232.
- Whanger P (2004). Selenium and its relationship to cancer: an update. *Br J Nutr* 91(01):11-28.
- Yu J, Hu S, Wang J, Wong GKS, Li S, Liu B, Deng Y, Dai L, Zhou Y, et al. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). *Science* 296(5565):79-92.
- Zhang L, Byrne PF, Pilon-Smits EA (2006). Mapping quantitative trait loci associated with selenate tolerance in *Arabidopsis thaliana*. *New Phytol* 170(1):33-42.
- Zuckerkindl E, Pauling L (1965). Evolutionary divergence and convergence in proteins. *Evolving Genes and Proteins* 97:166.