


RESEARCH ARTICLE

Open Access



Amino acid transporter (AAT) gene family in foxtail millet (*Setaria italica* L.): widespread family expansion, functional differentiation, roles in quality formation and response to abiotic stresses

Yang Yang^{1†}, Yongmao Chai^{1†}, Jiayi Liu¹, Jie Zheng¹, Zhangchen Zhao¹, Aduragbemi Amo¹, Chungue Cui¹, Qiumei Lu¹, Liang Chen^{1*} and Yin-Gang Hu^{1,2*} 

Abstract

Background: Amino acid transporters (AATs) play an essential role in growth and development of plants, including amino acid long-range transport, seed germination, quality formation, responsiveness to pathogenic bacteria and abiotic stress by modulating the transmembrane transfer of amino acids. In this study, we performed a genome-wide screening to analyze the AAT genes in foxtail millet (*Setaria italica* L.), especially those associated with quality formation and abiotic stress response.

Results: A total number of 94 AAT genes were identified and divided into 12 subfamilies by their sequence characteristics and phylogenetic relationship. A large number (58/94, 62%) of AAT genes in foxtail millet were expanded via gene duplication, involving 13 tandem and 12 segmental duplication events. Tandemly duplicated genes had a significant impact on their functional differentiation via sequence variation, structural variation and expression variation. Further comparison in multiple species showed that in addition to paralogous genes, the expression variations of the orthologous AAT genes also contributed to their functional differentiation. The transcriptomic comparison of two millet cultivars verified the direct contribution of the AAT genes such as *SiAAP1*, *SiAAP8*, and *SiAUX2* in the formation of grain quality. In addition, the qRT-PCR analysis suggested that several AAT genes continuously responded to diverse abiotic stresses, such as *SiATLb1*, *SIANT1*. Finally, combined with the previous studies and analysis on sequence characteristics and expression patterns of AAT genes, the possible functions of the foxtail millet AAT genes were predicted.

Conclusion: This study for the first time reported the evolutionary features, functional differentiation, roles in the quality formation and response to abiotic stresses of foxtail millet AAT gene family, thus providing a framework for further functional analysis of *SiAAT* genes, and also contributing to the applications of AAT genes in improving the quality and resistance to abiotic stresses of foxtail millet, and other cereal crops.

* Correspondence: chenliang9117@nwfau.edu.cn; huyingang@nwsuaf.edu.cn

[†]Yang Yang and Yongmao Chai contributed equally to this work.

¹State Key Laboratory of Crop Stress Biology for Arid Areas, College of Agronomy, Northwest A&F University, Yangling, Shaanxi, China
Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords: Foxtail millet, Amino acid transporter, Genome-wide characterization, Functional differentiation, Grain quality, Abiotic stresses

Background

Amino acids, as small-molecule organic compounds containing nitrogen, are the main carriers of nitrogen exchange in plants and the nutrients necessary for plant growth, development and quality [1]. The absorption and transport of amino acids are mainly mediated by amino acid transporters (AATs) [2]. Through the transmembrane transport of amino acids, amino acid transporters play essential roles in plant growth and development, including seed germination [3, 4], long-distance transport of amino acids [5], quality formation [6, 7], and response to pathogenic bacteria [8] and abiotic stress [9]. A total of 72, 65, 189, 85 and 297 amino acid transporters have been characterized in potato [10], *Arabidopsis* [1], soybean [11], rice [12] and wheat [13], respectively, suggesting that amino acid transporters are widely present in higher plants.

The AAT gene family in plants belongs to APC transporter superfamily, that comprises of two main families of amino acid/auxin permease (AAP) family and amino acid-polyamine-choline (APC) transporter family. The AAP family consists of Amino acid permeases (AAPs), lysine and histidine transporters (LHTs), auxin transporters (AUXs), *c*-aminobutyric acid transporters (GATs), neutral amino acid transporters (ANTs), proline transporters (ProTs), and amino acid transporter-like a/b (ATLa/b) proteins [8, 14, 15]. The APC family comprises cationic amino acid transporters (CATs), amino acid/choline transporters (ACTs), polyamine H⁺ – symporters (PHSs), and tyrosine-specific transporters (TTPs) subfamilies [10, 16].

Based on the heterologous expression system, the functions of AAT genes in *Arabidopsis* have been characterized in detail [17]. There are 8 AAP genes (*AtAAP1-AtAAP8*), and 6 *AtAAP* genes transport charged and neutral amino acids with various affinities [17, 18]. In endosperm, roots and cotyledon, *AtAAP1* is abundantly expressed, which is essential for the synthesis of storage proteins in *Arabidopsis* seed and seed yield, while regulating embryonic development and amino acid input in roots [14, 19]. *AtAAP2* and *AtAAP3* are mainly involved in the xylem-phloem transfer of amino acids and the absorption of amino acids in xylem, respectively [20, 21]. *AtAAP4* and *AtAAP5* can transfer amino acids to developing embryos and *AtAAP5* plays crucial role in the root amino acid intake system [22]. *AtAAP6* affects the interaction between plants and aphids through the regulation of the phloem's amino acid composition [8]. *AtAAP8* is mainly responsible for transporting amino acids into the endosperm and embryo at the early embryonic development

[23]. Additionally, the functions of some AAPs in other species have also been explored. In *Vicia faba*, *VfAAP1* is expressed abundantly in cotyledons and other pool tissues, and significantly affects the storage protein content (GPC) in seeds by regulating the nitrogen content of seeds [4, 6], while *VfAAP3* is significantly enriched in roots, stems, pods, pistils and seed coats at various stages of development [4]. In potato, *StAAP1* is expressed during the source-sink transition process, and the free amino acid contents are decreased in transgenic plants with reduced *StAAP1* expression [24], and the expression patterns of *StAAP4*, 5, and 6 may be associated with the long-distance transport of amino acids [10]. In barley, *HvAAP2* and 6 are mainly expressed in nutritive organs, while *HvAAP3* is high-expressed in seeds and contributes significantly to nitrogen accumulation in developing seed [25]. In rice, *OsAAP1* affects plant growth and grain yield by modulating the redistribution of neutral amino acids [26], and the quantitative trait locus qPC1 (QTL) for modulation of rice GPC is linked to the expression of *OsAAP6* [7]. In addition, several wheat *TaAAP* genes have been confirmed to co-localize with QTL of grain protein content [27].

In addition to the in-depth studies on AAP genes, the functions of some genes in other subfamilies have also been explored [5, 8, 9]. *AtLHT1* plays an indispensable role in the uptake of amino acids and mesophyll formation in root, while *AtLHT2* mainly involves in the movements and allocation of amino acids in floral organs [28–30]. *AtANT1* is expressed abundantly in flowers and stem leaves, transporting neutral and aromatic amino acids as well as indole-3-acetic acid (IAA) and arginine [31]. *AtAUX1* and *AtLAX3* affect root gravitropism and lateral root formation by facilitating IAA sink-source distribution and auxin-induced expression of root cell-wall-remodeling genes, respectively [32–35]. *AtProT1* and *AtProT2* participate in the long-distance transportation of proline and the absorption of proline and glycine betaine in roots of *Arabidopsis*, respectively [36, 37]. *AtProT3* is mainly expressed in leaf epidermal cells, which may be involved in the regulation of proline distribution in leaves [38]. *AtCAT5* mediates the reabsorption of amino acids permeated at leaf margin by transporting basic amino acids, while *AtCAT3*, *AtCAT6* and *AtCAT8* prefer to transport acidic or neutral amino acids [17]. *AtBAT1* has been identified as a bidirectional amino acid transporter expressed on mitochondrial membrane, which mediates γ -aminobutyric acid (GABA) from cytosol to mitochondria [39].

In addition to the role of *AAT* genes in seed germination, long-distance amino acid transport, root gravitropism, the formation of lateral roots and reproductive organs, a number of *AAT* genes play an important role in reducing damage to plants under various abiotic stresses, specifically by promoting the transportation of stress-response compounds and compatible solutes such as proline, GABA, and betaine [40]. For example, the *ProT* genes rapidly distribute proline to minimize damage to plants caused by water shortage [41]. *AtProT2* is strongly induced by salt stress, and plays a role in nitrogen distribution under salt stress [42]. *GmProT1* and *GmProT2* are sensitive to stress in the roots and transgenic plants can transport more proline after spraying proline [43]. *AtGATI* is significantly induced by GABA enrichment in plants under mechanical damage, stress or sensitive environment [44]. *OsAAP6*, *OsAAP11* and *OsANT3* are strongly up-regulated under salt and drought stresses, to regulate the tolerance to stress of plants through the transport of stress compounds [12].

Foxtail millet (*Setaria italica* L.), a member of the Poaceae family, is one of the oldest cereal crops, domesticated in Northern China. It is rich in essential amino acids, fatty acids and minerals, which is of important significance to human health [45]. To date, the *AAT* gene family in foxtail millet has not been characterized. Given their essential roles in the growth and grain quality formation of foxtail millet, we carried out a detailed characterization of *AAT* genes in foxtail millet. The aims of this study are to: (a) Accurately identify and characterize *AAT* genes in foxtail millet based on the latest genome sequence; (b) Explore the roles of paralogous and orthologous *AAT* genes in the build-up and functional differentiation of *AAT* family in foxtail millet through comparative analysis of their gene structure, sequence feature and expression patterns; (c) Mine the key candidate *AAT* genes that affect grain quality formation of foxtail millet through the comparative transcriptome analysis of two genotypes; (d) Detect the response of *SiAAT* genes to diverse abiotic stresses by Quantitative Real-Time PCR; (e) Predict the possible functions of the *SiAAT* genes based on the previous studies, sequence composition and expression pattern. This work will provide a deeper insight and understanding of the *AAT* genes in foxtail millet. Also, important clues for their functional analysis and applications in improving quality and resistances to abiotic stresses will be provided.

Results

Identification of *AAT* gene family in foxtail millet

Firstly, 104 putative *AAT* transcripts were identified in foxtail millet based on local BLASP. However, 94

high-confidence non-redundant *AAT* genes were confirmed after screening the conservative domains by HMMER, CDD and Interpro databases and removing the different transcripts of the same gene and sequences containing incomplete conserved domains, which was basically at the same level as other gramineous species besides hexaploid wheat (Table 1; Additional file 5: Table S1). Except for tetraploid soybean, *AAT* gene family in foxtail millet was expanded than that in *Arabidopsis* and potato. Overall, the number of *AAT* genes in monocots was higher than that in dicots. The *AAT* genes identified in foxtail millet were renamed according to their chromosomal localizations and phylogenetic relationships with other species. The length of *AAT* proteins in foxtail millet ranged from 311aa to 984aa, with molecular weight (Mw) varying from 4.97 kD to 107.5 kD, and isoelectric point (pI) ranged from 4.94 to 9.99 (Additional file 5: Table S1). Different subfamilies of *AAT* genes showed abundant diversity in subcellular localization, with the subfamily members of AAP, LHT, GAT, AUX and ProT all located on plasma membrane, while those of TTP, ACT, ANT and ATLb all located on vacuole membrane. Some members of the same subfamily showed different subcellular localization, such as, different genes of the ATLa subfamily located on both membranes of plasma and vacuole, while that of the CAT and PHS subfamilies located on the three membranes of the plasma, vacuole and chloroplast.

The number of predicted TM regions in *SiAATs* varied in different subfamilies, which ranged from 7 to 14 (Fig. 1a; Additional file 5: Table S1). For example, all AUXs contained 10 TM regions, ATLas contained 10 or 11 TM regions, while CAT and PHS subfamilies ranged from 9 to 14 and 7 to 12, respectively.

Chromosomal distribution and duplication events of *AAT* genes in foxtail millet

Here, 93 of the 94 *SiAATs* were mapped to 9 chromosomes, the remaining one was mapped to an unassembled scaffold, which were unevenly distributed, with 18 on chromosome 7 and 2 on chromosome 2 as the most and least, respectively (Additional file 1: Fig. S1). The apparent regional enrichment of *SiAATs* on some chromosomes was observed on their distributions. For instance, a large number of *SiAATs* were located on the ends of chromosomes 1, 5, and 7. Moreover, some *SiAATs* were mainly concentrated at the fronts of chromosomes 6 and 8 (Additional file 1: Fig. S1).

Of the 94 *AAT* genes in foxtail millet, 58 (62%) were involved in gene duplication events, including 36 tandem and 25 segmental duplications, of which *SiATLb12*, *SiLHT2* and *SiLHT9* were involved in both

Table 1 Comparison on the gene abundance of twelve subfamilies of *AAT* genes in 6 monocots and 3 eudicots

	Monocots						Eudicots		
	Millet ^a	Sorghum	Maize	Brachypodium	Rice	Wheat	Soybean	Arabidopsis	Potato
AAAP									
AAP	20(21.27%)	17(22.07%)	24(22.43%)	19(23.75%)	19(22.35%)	66(22.30%)	35(18.52%)	8(12.70%)	8(11.11%)
LHT	12(12.77%)	8(10.39%)	15(14.42%)	8(10.00%)	6(7.05%)	24(8.11%)	24(12.70%)	10(15.87%)	11(15.28%)
GAT	6(6.38%)	4(5.19%)	2(1.92%)	3(3.75%)	4(4.71%)	14(4.73%)	19(10.05%)	2(3.17%)	3(4.17%)
ProT	1(1.06%)	1(1.30%)	2(1.92%)	2(2.50%)	3(3.53%)	9(3.04%)	7(3.70%)	3(4.76%)	4(6.35%)
AUX	4(4.25%)	5(6.49%)	6(5.60%)	3(3.75%)	5(5.88%)	15(5.07%)	16(8.47%)	4(6.35%)	5(7.94%)
ATLa	6(6.38%)	6(7.79%)	7(6.54%)	6(7.50%)	7(8.24%)	18(6.08%)	16(8.47%)	5(7.94%)	8(11.11%)
ANT	2(2.13%)	4(5.19%)	3(2.8%)	5(6.25%)	4(4.71%)	18(6.08%)	6(3.17%)	4(6.34%)	5(7.94%)
ATLb	14(14.89%)	8(10.39%)	17(15.89%)	7(8.75%)	10(11.76%)	40(13.51%)	30(15.87%)	10(15.87%)	8(11.11%)
APC									
ACT	8(8.51%)	5(6.49%)	7(6.54%)	6(7.50%)	7(8.24%)	21(7.09%)	7(3.70%)	1(1.59%)	1(1.39%)
CAT	12(12.77%)	11(14.29%)	14(13.08%)	11(13.75%)	11(12.94%)	31(10.47%)	19(10.05%)	9(14.29%)	9(12.5%)
PHS	8(8.51%)	7(9.09%)	7(6.54%)	7(8.75%)	9(10.59%)	31(10.47%)	9(4.76%)	5(7.94%)	8(11.11%)
TTP	1(1.06%)	1(1.30%)	3(2.80%)	3(3.75%)	0(0.00%)	9(3.04%)	1(0.53%)	2(3.17%)	2(2.78%)
Total	94	77	107	80	85	296	189	63	72

^aThe numbers represent the number of identified *AAT* subfamily members, and the percentages in the brackets represent the proportion to all *AAT* genes

uplications (Additional file 1: Fig. S1; Additional file 5: Table S1). The 36 tandem duplications were classified into 13 groups, of which 2 groups with 5 genes (TD10, TD11), 1 group with 4 genes (TD3), 2 groups with 3 genes (TD1, TD13), and the remaining 8 groups all with 2 genes (Additional file 5: Table S1).

The 25 segmental duplications were classified into 12 groups, except for SD3 containing 3 genes, with other possessing 2 genes. Gene duplication events produced a large number of paralogous *AAT* genes, thus promoting the significant build-up of the *AAT* gene family in foxtail millet.

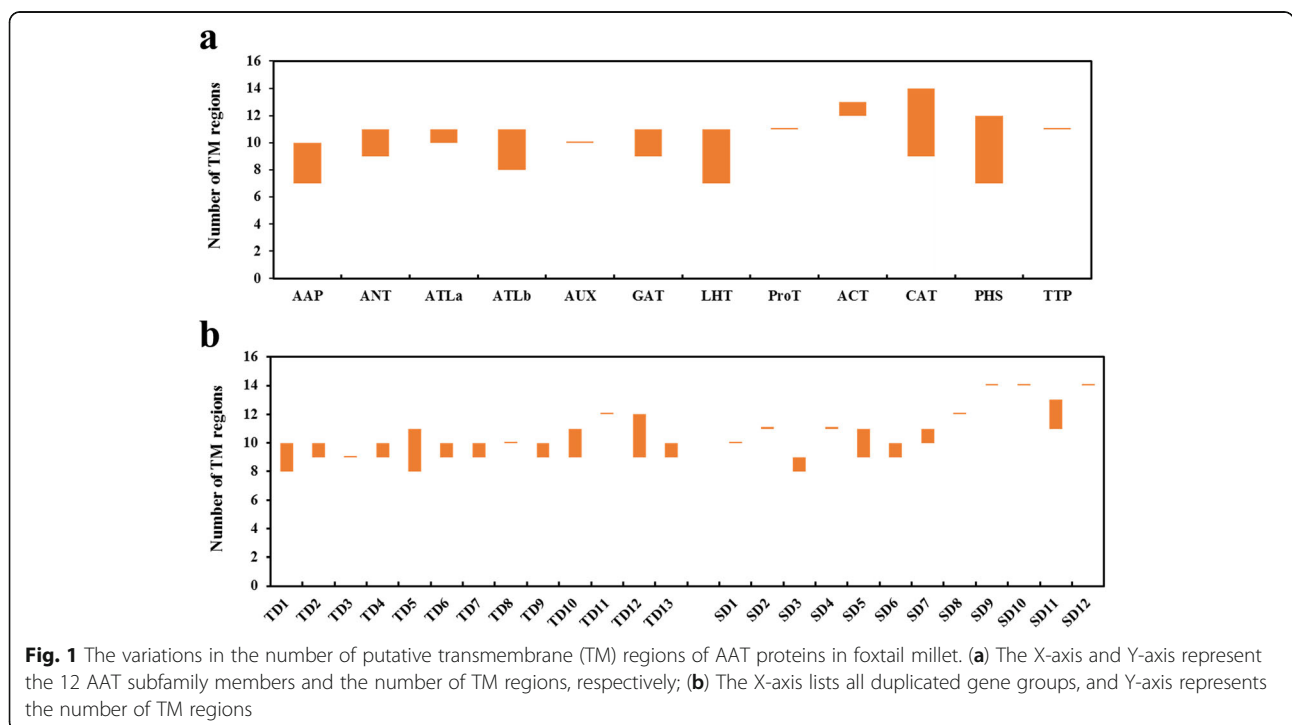


Fig. 1 The variations in the number of putative transmembrane (TM) regions of *AAT* proteins in foxtail millet. (a) The X-axis and Y-axis represent the 12 *AAT* subfamily members and the number of TM regions, respectively; (b) The X-axis lists all duplicated gene groups, and Y-axis represents the number of TM regions

The phylogenetic tree and structure of AAT genes in foxtail millet

The phylogenetic analysis with the AAT protein sequences from foxtail millet (94), rice (85), *Arabidopsis* (63) and potato (72) showed that they could be divided into twelve genetic groups with high confidence and that from Monocots and Eudicots distributed in the same group (Fig. 2). The 65 SiAAT proteins of the AAP family were divided into 8 subfamilies of amino acid permeases (AAPs, 20), lysine, histidine transporters (LHTs, 12), GABA transporters (GATs), proline transporters (ProTs, 1), aux transporters (AUXs, 4), amino acid transporter-like a (ATLa, 6), aromatic and neutral amino acid transporters (ANTs, 2) and amino acid transporter-like b (ATLb, 14). The 29 SiAAT proteins of the APC family were clearly distinguished into 4 subfamilies of cationic amino acid transporters (CATs, 12), amino acid/choline transporters (ACTs, 8), tyrosine-specific transporters (TTPs, 1), and polyamine H⁺-symporters (PHSs, 8) (Fig. 2).

Based on their annotation information, similar exon/intron structures were observed in most of the *SiAATs* in the same subfamily, such as *SiATLa1* and *SiATLa5*, *SiCAT3* and *SiCAT10*, and *SiATLb3* and *SiATLb2*. Variants were also found in some *SiAATs* of the same

subfamily, such as *SiAUX2* and *SiAUX3*, *SiAAP11* and *SiAAP18*, and *SiANT1* and *SiANT2* (Fig. 3). The conserved motifs of SiAAT proteins predicted with MEME were highly consistent with the phylogenetic relationship and classification of SiAATs (Fig. 3; Additional file 2: Fig. S2). Similar to their gene structures, the presence of the conserved motifs in different subfamilies of SiAAT proteins varied. For instance, Motif 1, 4, and 7 were widespread in the AAP family, while Motif 9 existed in the APC family. Some motifs were unique to certain subfamilies, for example, AUX subfamily only possessed Motif 19, and Motif 16 and 18 were only found in the ACT subfamily.

Variations in gene structure, conserved domains and sequences of paralogous *SiAAT* genes

The diversification of paralogous gene is one of the important sources of the functional differentiation of gene families [46–48]. The comparison on the TM regions of SiAATs found that significant variations in the number of TM regions occurred in the same duplicated gene groups, 60% of the duplicated gene groups (15/25) varied in the number of TM regions, which was 77% (10/13) and 42% (5/12) for the tandemly and for the segmentally duplicated gene groups, respectively (Fig. 1b; Table 2).

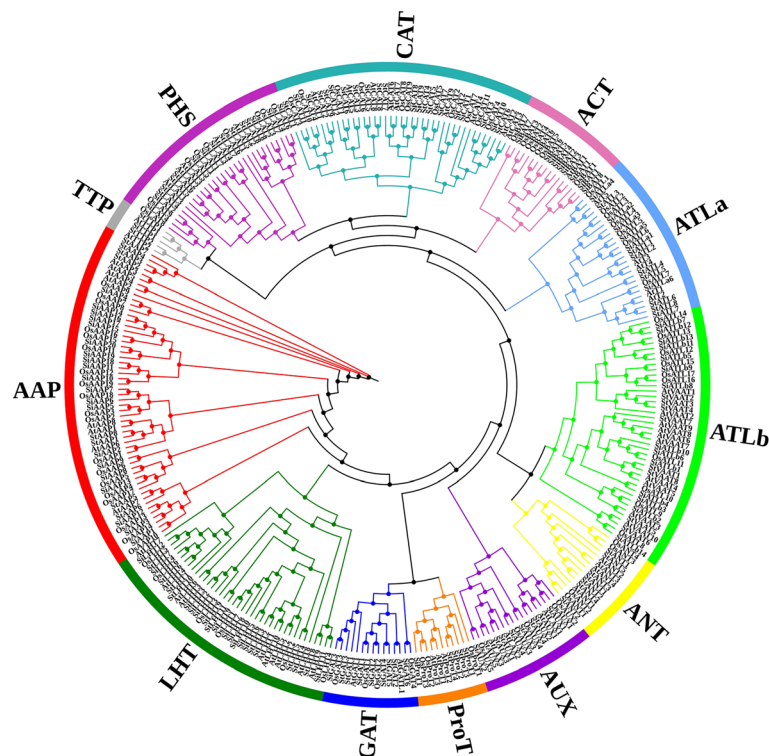


Fig. 2 Phylogenetic tree constructed by the AAT proteins of foxtail millet, rice, *Arabidopsis* and potato. MUSCLE and MEGA 6.06 were used to perform the multiple sequence alignment and to construct the phylogenetic tree with the maximum likelihood (ML) method and the JTT model. Different AAT subfamily members were shown with different colored branches. AAT proteins were named according to the ID in published articles, such as OsAAP1, AtAAP1 and StAAP1

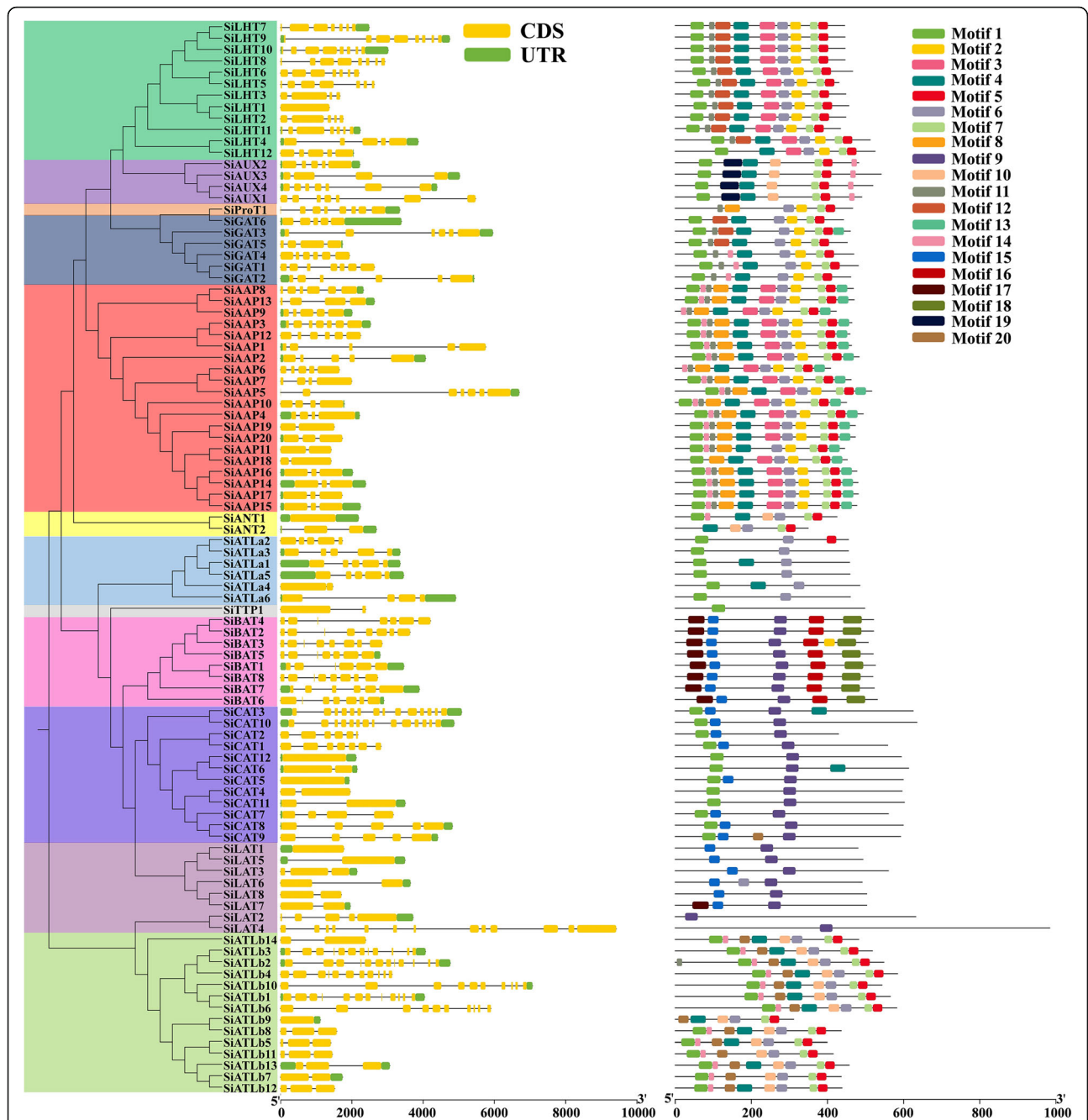


Fig. 3 Gene structure and conserved motifs of *SiAATs* in each subfamily. Different subfamilies of foxtail millet AAT family are shown with various color backgrounds. The gene structure was referred from the gene model file of foxtail millet reference genome. The exons, UTR regions and introns are displayed by green boxes, yellow boxes and the block lines, respectively. The detail motif sequences are as listed in Additional file 2: Fig. S2

Multi-sequence alignment of SiAAP proteins revealed that their overall similarity was 59.65% with 11 conserved motifs, and their TM regions were highly correlated with the conserved motifs, both in length and amino acid composition (Fig. 4). Motif 1 and 14 both formed the TM 1 and TM 2 regions, TM 4 and part of TM 5 regions included Motif 4. TM regions of 3, 6, 7, 8,

9 and 10 were with Motifs of 8, 6, 2, 7, 5, 13, respectively (Fig. 4). In addition, some conserved motifs were located in the non TM region, such as Motif 11 in the extra-membrane region and Motif 3 in the intra-membrane region (Fig. 4). It was worth noting that some SiAAPs had missing transmembrane regions due to incomplete conserved motifs, such as SiAAP6 and SiAAP9.

Table 2 The extensive variations in gene structure, protein structure and sequence of paralogous *AAT* genes produced by gene duplication in foxtail millet

Type of duplication	Subfamily	No. of duplicated gene groups	Gene and protein structure variation			Range of Ka/Ks values	Duplicated gene groups
			No. of TM variation	No. of gene structure variation	No. of conserved domain variation		
Tandem duplication	AAP	4	3	3	2	0.12–0.62	TD1-TD4
	ATLb	3	3	1	2	0.27–0.47	TD5-TD7
	GAT	1	0	1	1	0.43	TD8
	LHT	2	2	2	0	0.10–0.39	TD9,TD10
	ACT	1	0	1	1	0.18–0.27	TD11
	CAT	1	1	1	0	0.97	TD12
	PHS	1	1	0	1	0.41–0.67	TD13
	Sum ^a	13	10 (77%)	9 (69%)	7 (54%)	0.10–0.97 (0.28)	
Segmental duplication	ATLa	2	0	0	2	0.10–0.12	SD1,SD2
	ATLb	3	2	2	2	0.12–0.40	SD3-SD5
	LHT	2	2	1	0	0.12–0.37	SD6,SD7
	ACT	1	0	1	0	0.10	SD8
	CAT	4	1	1	3	0.14–0.50	SD9-SD12
	Sum ^a	12	5 (42%)	5 (42%)	7 (58%)	0.10–0.50 (0.21)	

^aThe data in columns 4, 5 and 6 represents the total number and its proportion (in the brackets) of mutated duplicated gene groups, while that in column 7 represents the range and average (in the brackets) of Ka/Ks values

Therefore, the number variation of TM regions was mainly determined by the presence of different conserved domains. Prediction of the secondary structure of SiAAT proteins found 14 α -helixes and 4 η -helixes structures, with vast majority located in the TM regions to ensure the efficient and stable transmembrane transport of amino acids (Fig. 4).

In terms of gene structure, 56% (14/25) of duplicated gene groups had variation in the number of introns, which was 69% (9/13) of the tandemly ones and 42% of the segmentally ones (Table 2; Additional file 3: Fig. S3). In addition, the variations in the conserved motifs were also observed in 14 duplicated gene groups, 7 for each of the two types of duplication (Table 2).

The Ka/Ks values of all paralogous *SiAAT* gene pairs were less than 1, ranged from 0.1 to 0.97, which suggested that these genes were subject to different levels of purifying selection (Fig. 5). The tandemly duplicated gene pairs had higher Ka/Ks values than the segmental duplications, with the average Ka/Ks values of 0.28 and 0.21, respectively (Table 2). The Ka/Ks values of tandemly duplicated genes of PHS subfamily ranged from 0.41 to 0.67, while that of PHS subfamily only ranged from 0.18 to 0.27. The Ka/Ks values of different subfamilies also showed different degrees of dispersion. For instance, that of *SiAAPs* ranged from 0.12 to 0.62, while that of *SiBATs* (ACT subfamily) from 0.18 to 0.27. This similar phenomenon was also observed on the Ka/Ks

values of segmentally duplicated genes (Fig. 5). In general, significant variations in gene structure, conserved domains and sequences among the paralogous *SiAAT* gene groups might greatly promote the functional diversity of the *AAT* gene family in foxtail millet.

Variations in expression levels of paralogous and orthologous *SiAAT* genes

The pattern of expression of all paralogous *AAT* genes were investigated using the public transcriptome data. Compared with the ancestral genes, according to their expression patterns, the newly duplicated genes could be classified into three main types of new-functionalization, sub-functionalization and non-functionalization, which were all observed in the tandemly duplicated gene groups with *SiAAP8/9* (TD2) possessing new functions. *SiAAP9* and *SiAAP8* were highly expressed during grain development, and in leaves at the filling stage, respectively. *SiAAP14/15/16/17* (TD3) and *SiCAT1/2* (TD12) showed sub-functionalization, as the expressions of *SiAAP16/17* and *SiCAT2* were down-regulated significantly. In addition, *SiATLb12* and *SiLHT3* losing their functions, as they were not expressed in any tissues (Fig. 6a). Though no new function was observed, similar functional differentiation were also observed in the segmentally duplicated gene groups, such as sub-functionalization of *SiATLb1/6/10* (SD3), *SiBAT1/SiBAT8* (SD8) and *SiCAT3/10* (SD9), and non-

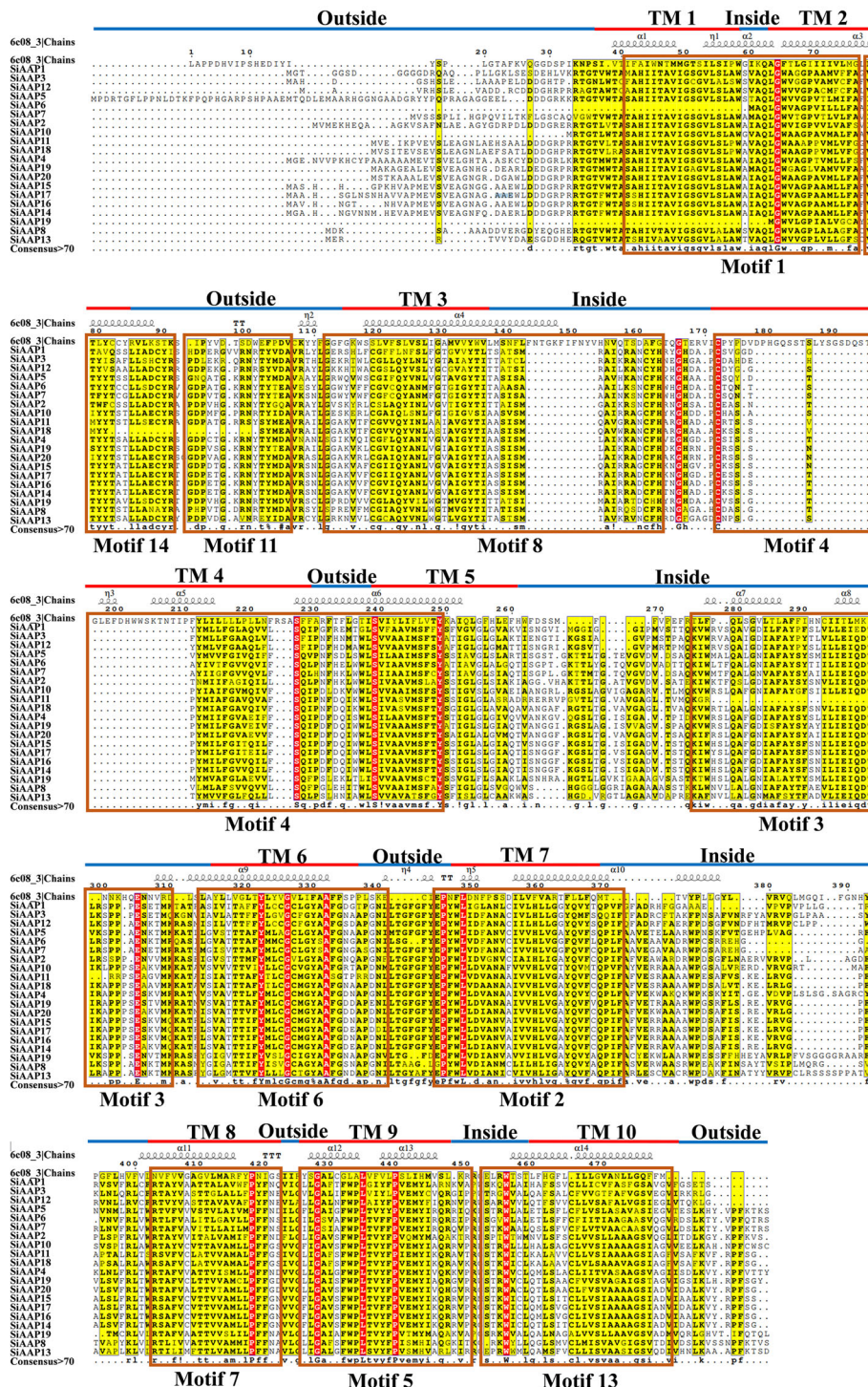


Fig. 4 The multiple-sequence alignment of the AAP proteins in foxtail millet. The TM regions of SiAAP proteins are indicated with red line, and their conserved motifs predicted by MEME are shown in the red boxes

functionalization of *SiATLa1/5* (SD1), *SiLHT1/2* (SD6) and *SiCAT4/11* (SD10) (Fig. 6b).

There were 62, 51, 52 and 35 orthologous AAT genes found in sorghum, rice, wheat and *Arabidopsis*,

respectively, and their corresponding relationships with foxtail millet were investigated (Fig. 7; Additional file 6: Table S2). The wheat genome was much larger than other species, so it was not displayed in Fig. 7. The

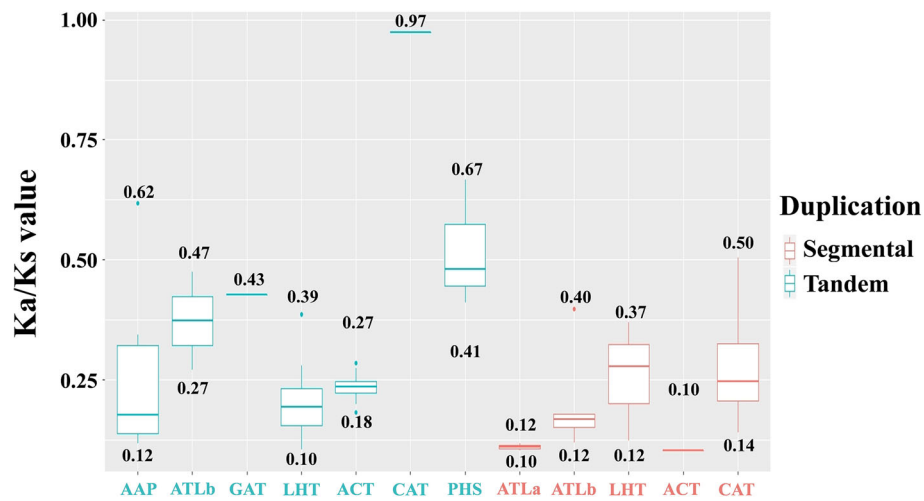


Fig. 5 The Ka/Ks values of homologous AAT gene groups in different AAT subfamilies of foxtail millet. The X-axis and Y-axis indicates the duplicated AAT gene pairs in different subfamilies and Ka/Ks (nonsynonymous / synonymous) values, respectively. Segmental duplications and tandem duplications were shown in red and blue, respectively

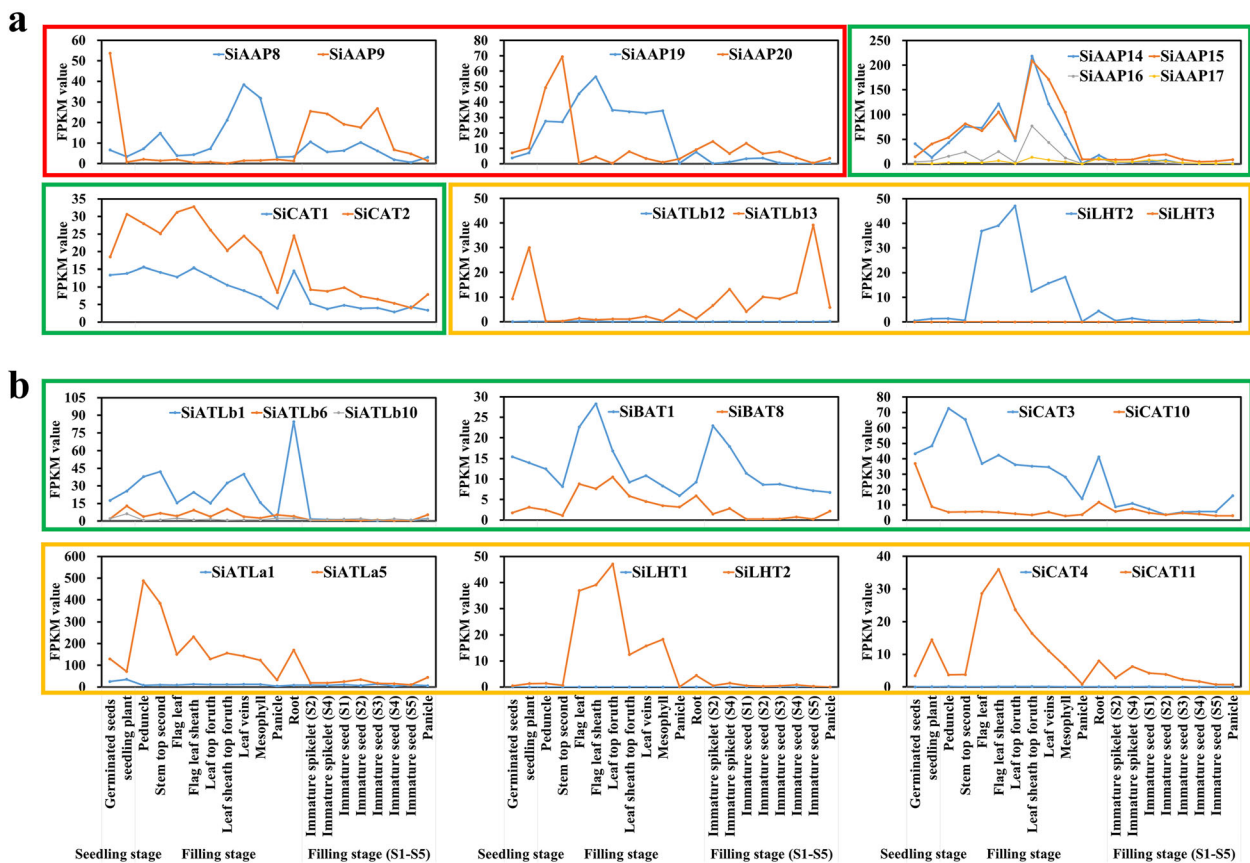
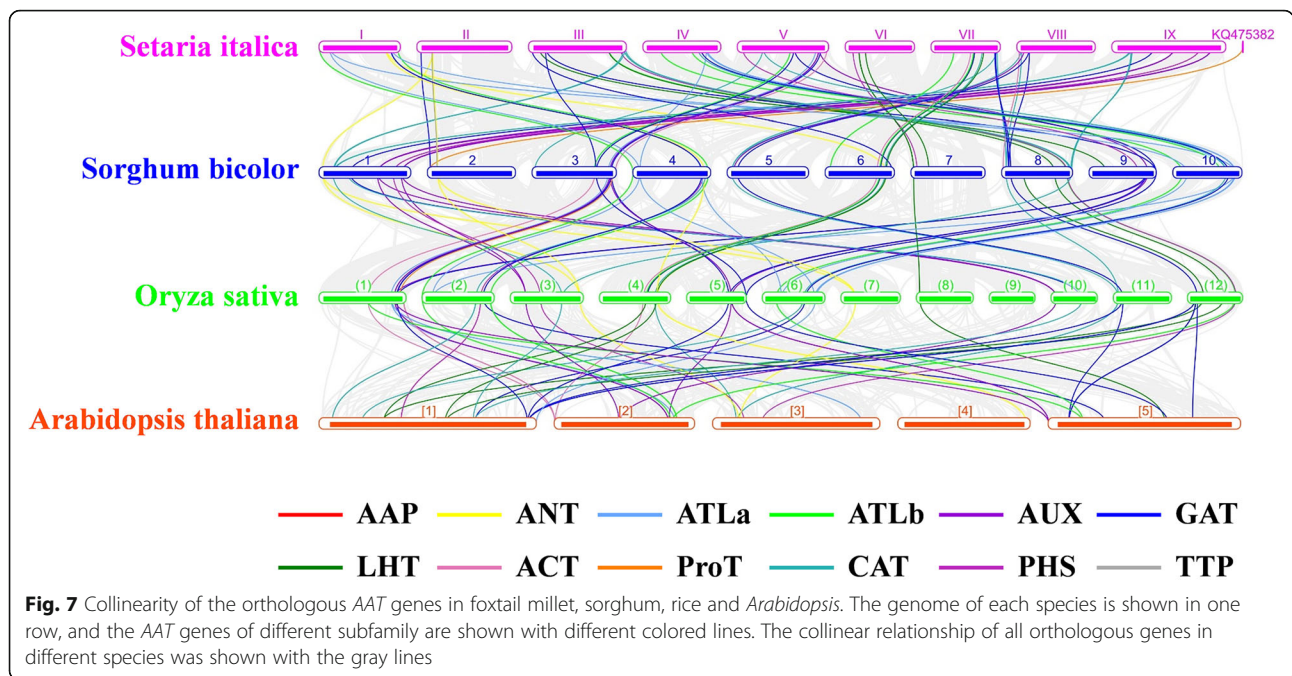


Fig. 6 Three variations of expression patterns of duplicated AAT gene groups in foxtail millet. (a) and (b) is for the tandemly and segmentally duplicated gene groups, respectively. The X-axis indicates different tissue at different stage, and the Y-axis represents the FPKM value. Different colored boxes represent different expression patterns. Red boxes: new-functionalization; Blue boxes: sub-functionalization; Yellow boxes: non-functionalization



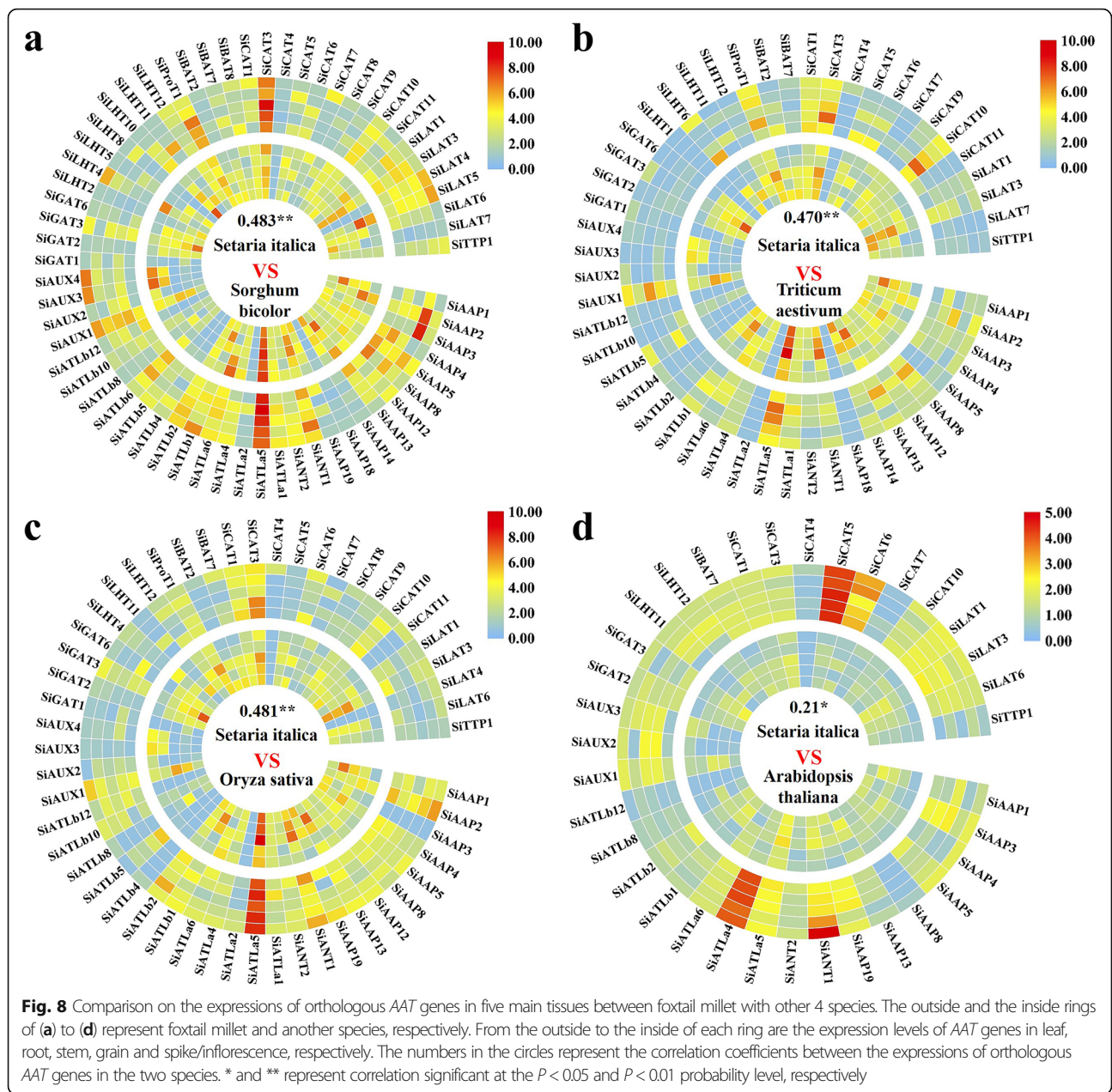
collinear relationship between different species was clearly distinguishable and showed good collinearity between chromosome 9 of foxtail millet and chromosome 1 of sorghum, chromosomes 3, 4, 5 of foxtail millet and chromosomes 8, 9, 10 of sorghum, respectively, etc. Moreover, the collinearity of the orthologous AAT genes among different species was consistent with that of the whole genomes. The number of orthologous subfamilies *LHTs* and *LATs* in millet and sorghum were significantly more than those in wheat and rice, which suggested that their expansion in millet and sorghum completed later than that in sorghum and wheat (Fig. 8; Additional file 6: Table S2). Compared with *Arabidopsis*, the AAP subfamily in grass species was significantly expanded. For instance, the orthologous of *SiAAP2*, *SiAAP12* and *SiAAP14* only could be found in grass species, which suggested that they might be produced after the differentiation of monocotyledon and dicotyledon.

The expression characteristics of those orthologous AAT genes in root, stem, leaf, inflorescence/spike and grain of foxtail millet, sorghum, wheat, rice and *Arabidopsis* were compared using their intergrated transcriptome datasets. Owing to the heterohexaploid nature of wheat with almost three homologous copies for each gene, the average expression values of these copies were used. The expression patterns of *SiAATs* in foxtail millet and their orthologous in other species were generally conservative, especially in the grass species (Fig. 8). The correlation coefficients on the expression patterns of AAT genes in gramineous species were 0.483, 0.470, and 0.481 between foxtail millet with sorghum, wheat and rice, respectively, while it was only 0.21 between foxtail

millet and *Arabidopsis*. Several high-expressed AAT genes in foxtail millet were also highly conserved in other species, such as *SiCAT3* and *SiATLa5*. In addition to the relatively conservative expression patterns of some important orthologous AAT genes among different species, the expression patterns of some genes were differentiated, such as *SiATLa6* and *SiATLb6* (Fig. 8). Compared with the orthologous genes in *Arabidopsis*, *SiAATs* showed greater variation in expression patterns, and several genes changed their tissue specificities, such as *SiATLb1* and *SiAUX3*.

Spatiotemporal expression patterns of *SiAATs*

The expression patterns of *SiAATs* were analyzed using the RNA-seq data of different tissues at multiple growth stages collected from the public online database. The heat map was displayed with the normalized log₂ (FPKM+ 1) values (Fig. 9a). According to the expression patterns, the *SiAATs* were clustered into three groups. The 8 *SiAAT* genes in the first group included 7 AAP family genes and one APC family gene with relatively high expression level, and were stably expressed in almost all tissues at different developmental stages. The 37 *SiAAT* genes in the second group were expressed in low abundance in most tissues, but expressed explicitly in some tissues, such as *SiAAP9* and *SiAAP7* were highly expressed in germinating seeds, while *SiATLb3*, *SiAUX3* and *SiAUX4* were relatively high expressed in panicle. The 49 *SiAAT* genes in the third group showed diverse spatiotemporal expression characteristics, such as *SiATLa6* and *SiATLa3* were highly expressed in the leaves at the seedling stage, but lower in the leaves at filling



stage; *SiAAP20* was specifically high-expressed in the stem, while *SiAAP3*, *SiBAT7*, *SiCAT11*, *SiBAT2*, *SiLAT7* and *SiLAT8* were expressed abundantly in leaf tissues including leaf sheath and mesophyll. In addition, some genes were highly expressed in two or more tissues, such as *SiAAP1*, *SiAAP13* and *SiAUX1* in stem and root; *SiANT1*, *SiCAT3* and *SiATLb1* were expressed abundantly in organs involved in the entire source-sink circulation including root, stem, leaf and grain. The expressions of the 20 selected *SiAATs* by qRT-PCR analysis were not completely consistent with that from the transcriptome data, but their expression characteristics and trends were similar, which verified

the reliability of the transcriptome data (Fig. 9a; Additional file 4: Fig. S4).

The transcriptome analysis of spike RNA-seq data showed that about half of *SiAATs* were expressed lowly during seed development, while the remaining ones were expressed at relatively high levels at different stages (Fig. 9b). *SiAAP2*, *SiAAP9*, *SiATLa5*, *SiLAT5*, *SiATLb2* and *SiBAT1* etc, were highly expressed in spikelets and grains throughout the grain development process (from S1 to S5). *SiATLa6*, *SiAAP8*, *SiAAP20*, *SiAAP1*, *SiLAT6*, *SiCAT2* and *SiCAT3*, etc were highly-expressed at the early stage of grain development. *SiGAT6* and *SiProT1*, etc were highly expressed at specific stages. In addition,

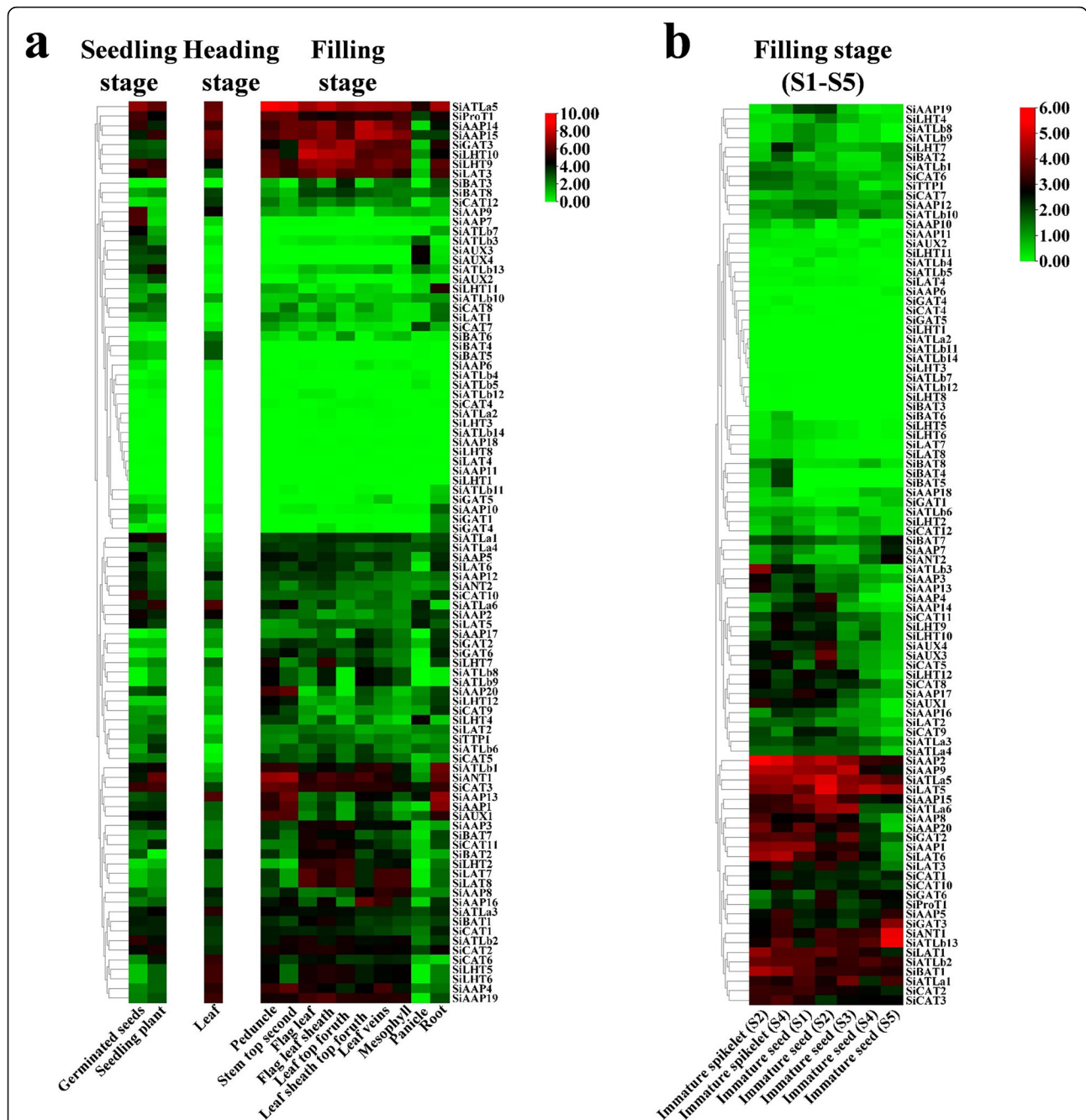
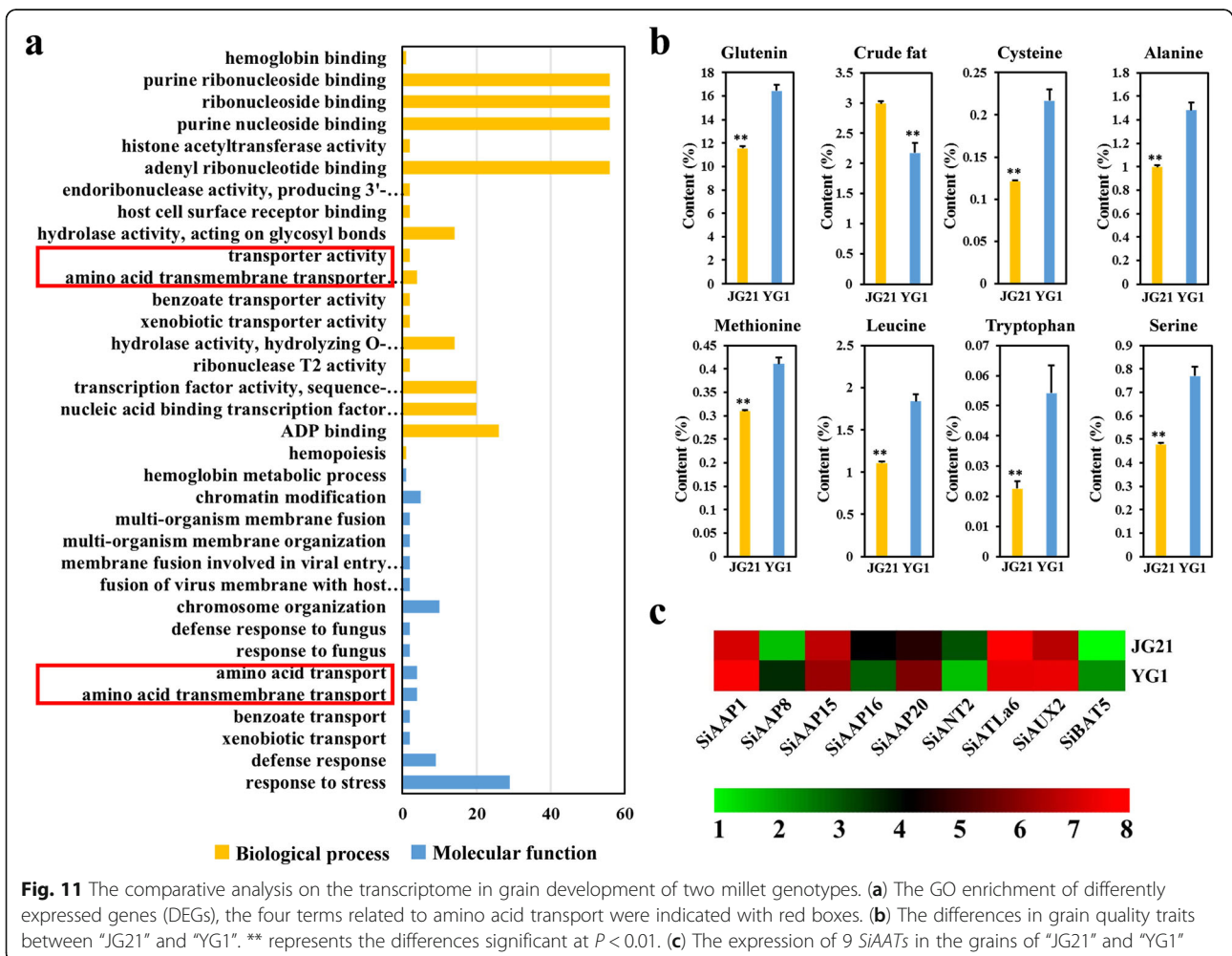
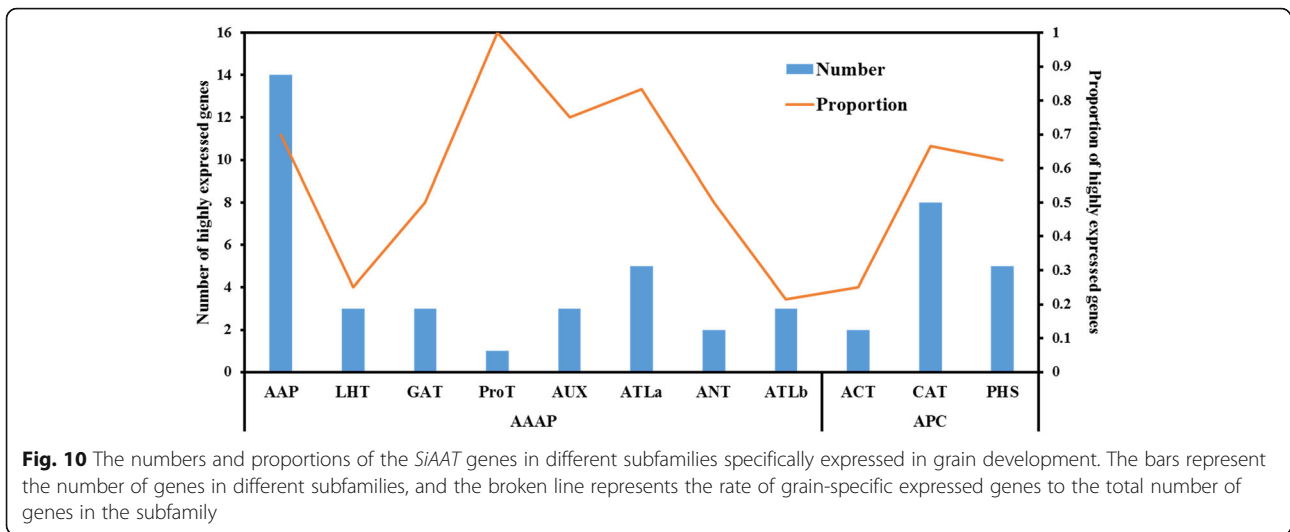
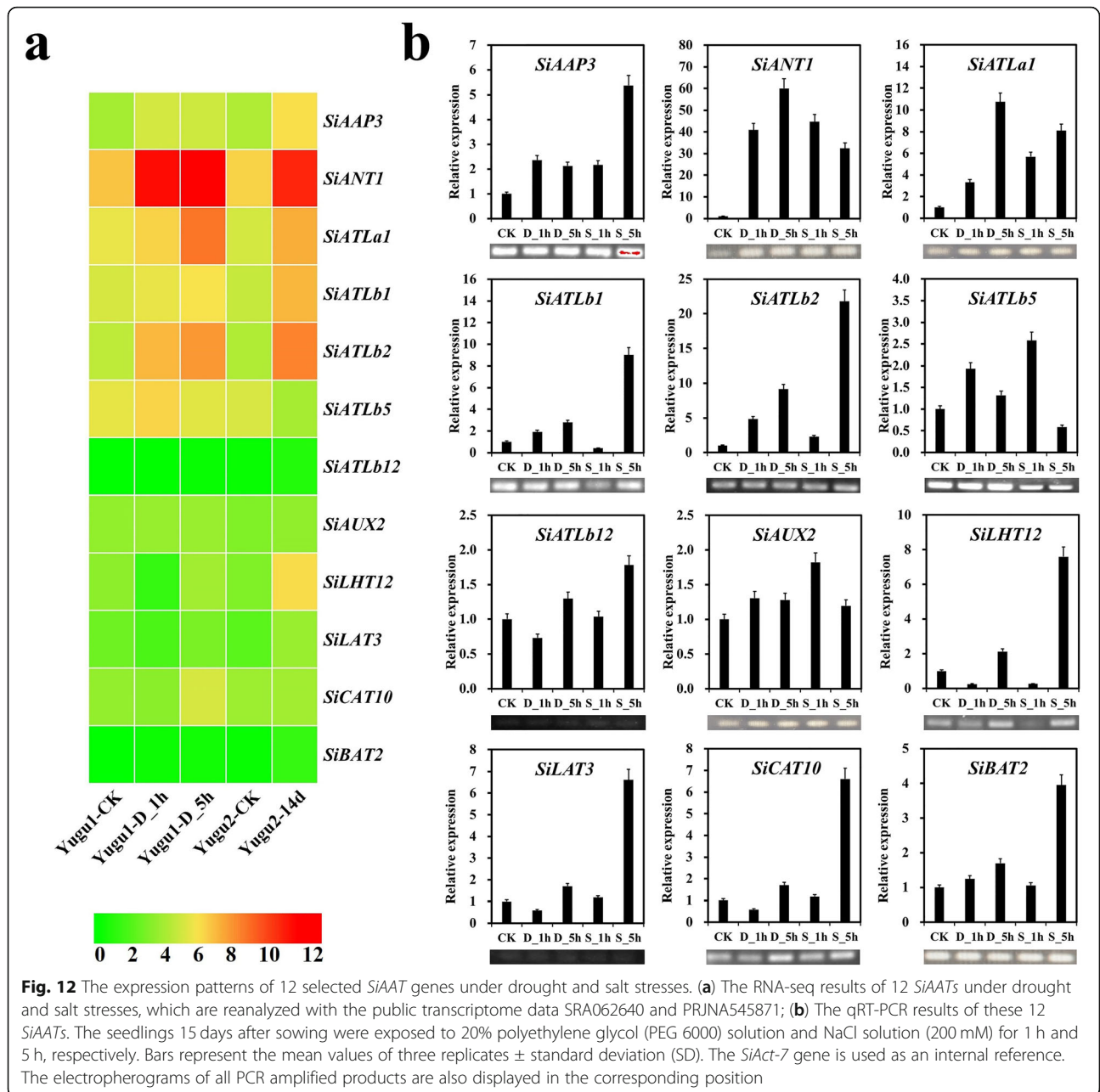


Fig. 9 Expression patterns of *SiAATs* in multiple tissues and developing grain of foxtail millet. (a) The expression patterns of *SiAAT* genes in 13 tissues at three growth stages, which were reproduced based on the transcriptome data; (b) The expression pattern of *SiAATs* in developing grain, which were obtained from RNA-seq. Tissue name is listed at the bottom of each column, while the color scale is shown at the right. The $\log_2(\text{FPKM} + 1)$ values are used to display the expression levels

the numbers of the highly expressed *SiAATs* during grain development in different subfamilies were counted, as 14 *SiAAPs*, 8 *SiCATs*, 5 *SiLATs*, 5 *SiATLAs* and 3 *SiAUXs*, which accounted for 70% of total *SiAAPs* (20), 67% of *SiCATs* (12), 63% of *SiBATs* (PHS, 8), 83% of *SiATLAs* (6) and 75% of *SiAUXs* (4), respectively (Fig. 10).

There were significant differences in multiple grain quality traits between the two foxtail millet genotypes “JG21” and “YG1”. Except for crude fat content, the basic amino acid contents of glutenin, cysteine, alanine, methionine, leucine, tryptophan and serine in “JG21” grain was significantly lower than “YG1” (Fig. 11b). The developing grains (including glumes) of “JG21” and “YG1” at filling





stage in 2019 were separated for RNA extraction and RNA-seq. Finally, 348 differentially expressed genes (DEGs) in the developing grains of these two genotypes were identified, of which 164 and 184 DEGs were down-regulated and up-regulated in “JG21”, respectively. GO analysis showed that DEGs were mainly enriched in multiple terms such as responses to fungus or viruses, ribonuclease activity, stress response, and transmembrane transport of various substances (Fig. 11a). Interestingly, 9 *AAT* genes were enriched into four terms related to amino acid transport, including 5 *SiAAPs*, 1 *SiANT*, 1 *SiATLb*, 1 *SiAUX* and 1 *SiBAT* (Fig. 11c). The differential expression

of these *SiAATs* suggested that these genes might directly affect the formation of grain quality.

Response of *SiAATs* to abiotic stresses

The qRT-PCR results of JG21 and the transcriptome data of the foxtail millet cultivars of Yugu1 and Yugu2 showed a relatively consistent trend of a similar response of *SiAAT* genes to abiotic stress (Fig. 12a). The qRT-PCR analysis of 12 *SiAATs* in foxtail millet seedlings (15 days after sowing) to simulate drought (20% PEG 6000, 1 h and 5 h) and salt (200 mM NaCl, 1 h and 5 h) stress revealed that various subfamilies of *SiAATs* responded differently to abiotic stresses

(Fig. 12b). Seven *SiAAT* genes were up-regulated more than 5 times after drought or salt stress, of which *SiANTI* was up-regulated by about 60 times at 5 h after drought stress. *SiANTI*, *SiCAT10* and *SiATLa1* were mainly induced by drought stress, and their expression all reached the highest levels at 5 h after drought treatment. Except for *SiAUX2* and *SiATLb5*, the other 7 *SiAAT* genes reached their highest expression levels at the late stage of salt stress (5 h). *SiLHT12*, *SiATLb1* and *SiLAT3* were specifically and strongly induced by long-time salt stress, while *SiANTI*, *SiATLa1* and *SiATLb5* had similar expression patterns under both drought and salt stress.

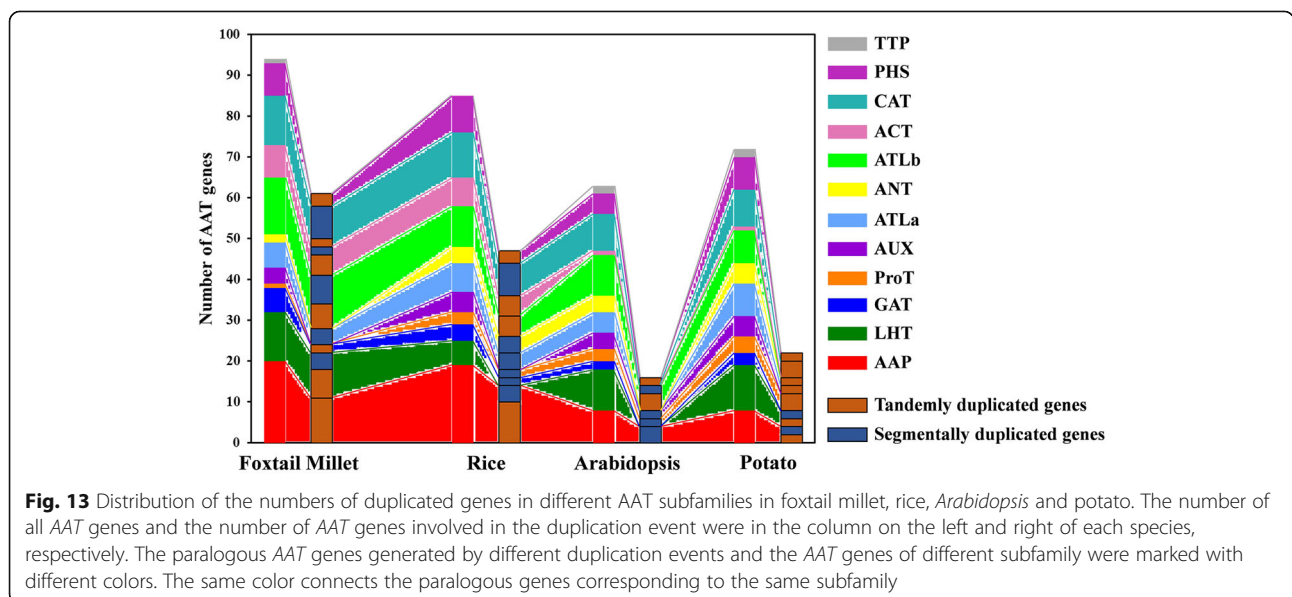
The response of different *SiAATs* to abiotic stress also showed temporal specificity. *SiAAP3* was up-regulated under 1 h drought stress, while it was up-regulated at 5 h under salt stress. *SiLHT12* was down-regulated under 1 h drought and salt stresses, but then was strongly up-regulated at 5 h (Fig. 12b). *SiANTI* was continually up-regulated with an increase of drought stress, and was down-regulated under 5 h salt stress. *SiATLb5* and *SiATLa1* responded at the early and late stages of drought and salt stress, respectively. Furthermore, some *SiAATs* showed a sustainable response to abiotic stress; for example, both *SiATLb2* and *SiATLa1* were continually up-regulated under both salt and drought stress. These genes might enhance the adaptability of foxtail millet to abiotic stresses through their active responsiveness.

Discussion

A large number of tandem and segmental duplication events were the direct cause of AAT gene family expansion in foxtail millet

Amino acid transporters (AATs) play indispensable roles in plant growth by transporting and distributing different types of amino acids, and therefore are

considered as important targets for crop improvement [15]. Here, a total of 94 *AAT* genes were characterized in foxtail millet, which were divided into 12 subfamilies and was consistent with the previous reports in other plants (Table 1). Different subfamilies of *AATs* in monocotyledons and dicotyledons were corresponding, thus confirming that they were formed before the differentiation in monocotyledons and dicotyledons (Fig. 2). The number of *SiAATs* in foxtail millet was less than those in hexaploid wheat (294) and maize (107) [13], but more than that in sorghum (77), rice (85) [12], and *Brachypodium* [13], and much higher than that in *Arabidopsis* (63) [1] and potato (72) [10]. Earlier studies indicated that the plant *AAT* gene family originated from the prokaryotic genome through horizontal gene transfer and subsequent duplication events [49]. There were 47, 16, 22 *AAT* genes associated with duplication events in rice, *Arabidopsis* and potato, accounting for 55.29% (47/85) [12], 25.40% (16/63) [10] and 30.55% (22/72) [10], respectively. As observed in this study, 58 of the 94 (62%) *SiAATs* were duplicated in foxtail millet, including 36 and 25 tandemly and segmentally duplicated genes, respectively. All these further verified the influence of gene duplication of the *AAT* family expansion in monocots. Furthermore, obvious imbalances in the expansion of different *AAT* subfamilies in monocots and eudicots were found, as AAP, ATLa/b, ACT and CAT subfamilies expanded greatly in monocots, while more *ProT* genes in eudicots (Table 2; Fig. 13). These increased *AATs* may play important roles in specific functions, and promoted further differentiation of monocots and eudicots.



The multi-levels variations of paralogous AAT genes generated from gene duplication promoted the functional diversity of AAT family in foxtail millet

Gene duplication and subsequent functional differentiation of new genes promoted the expansion and new-functionalization of important gene families, thereby enabling plants to better adapt to the environment [46]. Variations in the sequence, structure and expressions of genes were the three main sources of functional differentiation [13]. Base mutation was the most direct evidence of sequence variation. In general, the Ka/Ks values of all duplicate groups except for TD12 were far less than 1, indicating that most of the paralogous AAT genes undergone purifications in sequence variation to ensure the stability of key biological functions (Fig. 5). The Ka/Ks values of different AAT subfamilies showed different median and dispersion, such as AAP, PHS and CAT subfamilies had higher dispersion, which indicated that these subfamilies may have the opportunity to produce new features. However, the Ka/Ks values of ACT subfamily, both tandemly and segmentally duplicated gene pairs, showed smaller dispersion, suggesting that their functions might be more conservative. In addition, tandemly duplicated gene pairs showed higher Ka/Ks values than segmentally duplicated gene pairs, suggesting they were under stronger selection pressure, which confirmed that tandem duplication played more crucial roles in functional diversity of foxtail millet AAT genes. Similar results were also found in the study of wheat AAT family genes [13]. These results confirmed that the sequence variations of paralogous genes in different AAT subfamilies of foxtail millet increased the diversity of AAT genes.

Previous studies have shown that variations in gene structure are also important markers of gene family evolution [48, 49]. In this study, 56% (14/25) of the paralogous gene groups had significant changes in the number of introns, resulted in 9 and 5 groups of tandem and segmental duplications, respectively, accounting for 69 and 42% of the total (Table 2; Additional file 3: Fig. S3). The variations in the number of TM regions were found in 77% (10/13) and 42% (5/12) of the tandemly and segmentally duplicated gene groups, respectively, which were caused by the deletion of conserved domains, such as SiAAP5/6/7 (TD1), SiAAP8/9 (TD2) (Fig. 1b; Fig. 4). There were significant variations in gene structure and the number of TM regions which was observed in 56% (14/25) and 60% (15/25) of the two duplicated gene groups, confirming that the structural variation of paralogous genes was one of the important sources of the functional categorization of the AAT family genes in foxtail millet.

Changes in expressions of paralogous genes are another key sources of functional categorization of important gene family [49]. Sub-functionalization and new-

functionalization of new duplicated genes directly lead to significant differences in expression level and spatio-temporal expression patterns, compared to their ancestor genes [48]. As observed in rice [12], soybean [11] and potato [10], the changes of the three expression types of paralogous genes were also observed in this study. For instance, *SiAAP8/9* (TD2) and *SiAAP19/20* (TD4) were specifically expressed in different tissues, (Fig. 6), confirming that the expression changes of paralogous genes in AAT family were common in higher plants. Interestingly, the paralogous genes with new functions were only found in tandemly duplicated gene groups, suggesting that tandem duplication has a greater contribution to the evolution of the AAT gene family in foxtail millet.

The changes of orthologous AAT genes in foxtail millet play important roles in differentiation of new functions

Along with the separation of species and the influence of natural selection, though the orthologous genes of different species are largely conserved, the variations in their regulatory elements, epigenetic markers, *etc* will cause changes in the pattern of expression of orthologous genes in various species, resulting in their functional differentiation [50–54]. For some C_4 genes in C_4 species, such as *PEPC*, *MDH* and *PPDK*, their C_4 functions were directly obtained by the differences in the subcellular localization and expression level of orthologous genes [55]. The completion of genome sequencing allowed us to quickly identify the orthologous genes of different plants at genome level (Fig. 7). There were 62, 51, 52, 35 orthologous of the *SiAATs* identified in sorghum, rice, wheat and *Arabidopsis*, which were consistent with the genetic distance and differentiation time of these species [56]. The AAP subfamily of grass species expanded significantly after the differentiation of monocots and eudicots, which was consistent with previous studies [12, 13]. Within the grass family, the LHT and LAT subfamilies expanded significantly in foxtail millet and sorghum, which may have improved their adaptability to the environment. Most *SiAATs* and their orthologous in other species had similar expression patterns, especially for those of grass family with high correlation coefficients ranged from 0.470 to 0.483 (Fig. 8). Compared with their orthologous in *Arabidopsis*, the tissue expression characteristics of some *SiAATs* had changed, which may be related to the long genetic distance and complete differentiation between foxtail millet and *Arabidopsis*.

Many important AAT genes were expressed conservatively in multiple species to retain essential biological functions

Spatiotemporal expression analysis of important gene family members will help to predict their putative

functions [57]. Some *SiAATs*, such as *SiProT1* and *SiATLa5*, were abundantly expressed during the whole period of growth, which were essential for the growth and development of foxtail millet (Fig. 9). Some *SiAATs* were abundantly expressed in specific organs at certain developmental stages, such as *SiAAP9*, *SiAAP7* and *SiAAP2* in germinating seeds, *SiAAP1* and *SiAUX1* in stems and roots, and *SiLAT7* and *SiLAT8* in leaf tissues, while, *SiAAT6* was highly expressed only in leaves at seedling and heading stages, but low at filling stage. The different spatiotemporal expression patterns of these genes suggested that they might have specific functions.

The functions of some *AAT* genes have been extensively studied in the model species, such as *AtAAP1* for seed germination and amino acid intake in root [9, 14, 58]. It was found that *SiAAP7* and *SiAAP2*, *OsAAP1* [26] possessed similar functions; *AtAAP3* in the uptake of amino acids in xylem, and *SiAAP4* highly expressed in stems and veins [21]; *StAAP1* was expressed in multiple organs of source-sink transport in potato [10, 24]; *AtLHT1* located in root epidermis and mesophyll cells in *Arabidopsis*, which was primarily responsible for the amino acid uptake in soil [28, 29]; *AtAUX1*, *AtLAX3* and *OsAUX1* related to the gravitropism of root and lateral root formation [12, 32, 33]. As similar expression patterns were also observed for orthologous genes, which maintain the important process of plant growth and development. Therefore, we preliminary predicted their functions in foxtail millet based on these previous studies, the phylogenetic relationships and the spatiotemporal expression characteristics of *SiAATs* in foxtail millet (Table 3).

Multiple *AAT* genes may contribute directly to grain quality traits in foxtail millet

The *AAP* genes are closely related to grain development and quality formation in many species. *AtAAP1* in *Arabidopsis* and *OsAAP6* in rice are directly related to grain protein content [7, 14], multiple *AAT* genes were located in or near the QTL regions associated with quality traits in wheat by genome-wide association analysis (GWAS) [27]. The RNA-seq analysis of the foxtail millet grains at development stages revealed that 70% *AAPs* (14/20), 67% *CATs* (8/12), 63% *BATs* (5/8), 83% *ATLa* (5/6) and 75% *AUXs* (3/4) were highly expressed during grain development, including *SiAAP9* and *SiAAP8*, the orthologous genes of *OsAAP6*, suggested that they played a major role in the development and quality formation of foxtail millet grain (Fig. 10). The DEGs between the two cultivars with extremely significant differences in the content of glutenin, crude fat, four non-polar amino acids (Alanine, Methionine, Leucine and Tryptophan) and two neutral amino acids (Cysteine and Serine)

were enriched into four GO terms related to amino acid transport (Fig. 11a). Nine *SiAATs*, including *SiAAP8*, belonging to *AAP*, *ANT*, *ATLa*, *AUX* and *PHS* subfamilies, were expressed differentially in the developing grains of the two genotypes, which suggested that they might contribute to the grain quality traits (Fig. 11c). In *Arabidopsis*, *AtAAP1* and *AtAAP2* directly affected the grain protein content, and *AtAAP2* also significantly affected the fat content in the grain [20]. The differential expression of their homologs *SiAAP1* and *SiAAP8* in the two foxtail millet grains may lead to their difference in grain protein and crude fat content. *Arabidopsis AtAAP6* mutant showed a decrease in the content of lysine, phenylalanine, leucine and aspartic acid [8]. Therefore, it is speculated that the difference in leucine may be caused by the differential expression of *SiAAP20*, a homologous gene of *AtAAP6* in foxtail millet. In addition, neutral amino acid transporters (*ANTs*) could specifically transport aromatic and neutral amino acids. The content of such amino acids including cysteine, tryptophan and serine may be affected by the expression level of *SiANT2*. However, the functions of those differentially expressed *SiAAT* genes on the contents of different amino acids and protein content in foxtail millet grains should be verified in detail in further experiments. All these suggested that it might be feasible to improve the quality traits by manipulating the expressions of those important *SiAAT* genes in foxtail millet, and our results provided a base for further analysis of their function in quality formation in foxtail millet.

SiAATs enhanced the adaptability of foxtail millet to abiotic stresses

Drought and salt seriously affect crop yield and grain quality, and are the most important abiotic stresses in agricultural production. In *Arabidopsis* and other species, it has been confirmed that *AAT* genes in higher plants regulate the balance of osmotic potential to resist the impact of stress on plant growth, mainly through the transport of a variety of stress-response compounds and compatible solutes, such as betaine, GABA, and proline [40–44]. *AtAAP4* and *AtAAP6* were down-regulated under salt stress [42], and *OsAAP4*, *OsAAP8* and *OsBAT4* were down-regulated under both salt and drought stress [12]. *AtProT2* [42], *McAAT1* [60] and *HvProT* [61] were greatly triggered by salt stress. In this study, active responsiveness to simulated drought and salt stresses were observed in 12 *SiAAT* genes investigated, and various response patterns to different stresses were found in different subfamily members (Fig. 12). It should be noted that *SiAATs* located on vacuole membranes, such as *SiANT1* and *SiALTB2*, responded

Table 3 Functional prediction of AAT genes in foxtail millet based on homology relationship, expression pattern and previous studies

Subfamilies	AATs in foxtail millet	expression organization ^a	Possible biological processes	Previous studies
AAP	SiAAP7;	GS;	The intake of amino acid in endosperm and embryo development;	AtAAP8,1 [23, 58];
	SiAAP1,5,13;	R,ST;	Amino acid absorption in root and transport in stems;	AtAAP7,1 [19];
	SiAAP2,10,9;	GS,SP,G	Play a role in early embryo development and formation of floral organs and grain quality;	AtAAP8,1,VfAAP1,OsAAP6 [4, 6, 7];
	SiAAP3,8,12,14	LS,LV, G	Participate in the transport and distribution of amino acids in leaf and grain formation;	
	SiAAP4,15,19, 20	ST,LS,LV	Participate in the transfer and distribution of amino acids from stems to leaf	HvAAP3,AtAAP7 [25];AtAAP2,3,6,OsAAP8,15; HvAAP2,6 [8, 12, 20, 21]
GAT	SiGAT3;	LS,LV,L,SP	Maintain cell ion balance under normal conditions;	AtGAT1 [44]
	SiGAT6	R,ST,LS,SP	Amino acid absorption in root and long-distance transportation of GABA	
ProT	SiProT1	ST,L,GS	Participate in the long-distance transport of proline and regulate the distribution of proline in leaf	AtProT1,AtProT3 [5, 38]
LHT	SiLHT11	R;	Amino acid uptake in root;	AtLHT1,9 [28, 59]
	SiLHT2,5,10	L;	Transport and distribution of amino acids in leaf;	
	SiLHT9	LS,ST,R	Long-distance transportation of amino acids from root to leaf	
AUX	SiAUX1;	ST,R;	Participate in the geotropism of root and the transport of auxin in the stem;	AtAUX1 [32–34];
	SiAUX2,3,4;	SP; G	Participate in spikelet development and grain formation	OsAUX1,StAUX [10, 12]
ANT	SiANT1;	ST,R,LS,LV;	Participate in long-distance transportation of amino acids;	AtANT1 [31]
	SiANT2	SP,G,GS	Participate in seed germination, spike differentiation and grain formation	
ATLa	SiATLa1;	GS,SE;	Participate in seed germination;	NA
	SiATLa3,4,5;	ST,LS,LV,R,SP;	Participate in long-distance transportation of amino acids;	
	SiATLa6	L,SP	Regulate the distribution of amino acids in leaf and Participate in spike differentiation	
ATLb	SiATLb3,6,10, 13;	GS,SE,SP,G;	Participate in seed germination, spike differentiation and grain formation;	NA
	SiATLb1;	R,ST,LV;	Long-distance transportation of amino acids from root to leaf;	
	SiATLb7;	GS;	Participate in seed germination;	
	SiATLb8;	ST,LS,LV	Long-distance transportation of amino acids from stem to leaf;	
CAT	SiCAT10	GS	Participate in seed germination;	NA
	SiCAT3, 5, 6, 8, 9	GS, ST, R, L	Participate in long-distance transportation of amino acids;	
	SiCAT7	SP	Specific influence on spike differentiation;	
	SiCAT11, 12	SP, L	Participate in spikelet formation and amino acid transport and distribution in leaves	
ACT	SiBAT1, 2, 7	ST, L, SP	Participate in spikelet formation and long-distance transportation of amino acids	NA
PHS	SiLAT1, 2, 5	GS, SP, G;	Participate in seed germination, spike differentiation and grain formation;	NA
	SiLAT3	GS, LS, LV, ST, R	Participate in spikelet formation and long-distance transportation of amino acids	
TTP	SiTTP1	SE, SP	Affect seedling growth and spikelet differentiation	NA

^aGS Germinated seed, R Root, ST Stem, SP Spikelet, G Grain, SE Seedling, L Leaf, LV Leaf Vein, LS Leaf Sheath

stronger to these abiotic stress, which may relate to their roles in rapid adjustment of cell osmotic potential to maintain cell water balance. These genes could be used as key candidate genes to improve resistance and tolerance under abiotic stresses.

Conclusions

In this study, 94 *AAT* genes in foxtail millet were systematically identified and characterized, which were divided into 12 subfamilies. The paralogous genes generated by tandem duplication and segmental duplication promoted *AAT* gene family expansion. The variations in sequence, structure and the pattern of expression of tandemly duplicated paralogous *AAT* genes had significant impacts on their functional differentiation. The functional differentiation of *SiAATs* was the result of the unequal expansion and differentiation of paralogous genes in different subfamilies. The transcriptome analysis revealed the contributions of some *AAT* genes in the formation of millet quality. Multiple *SiAAT* genes actively responded to drought and salt stresses as revealed by qRT-PCR. The possible functions of the *SiAAT* family members were predicted. These findings might provide valuable information for further functional analysis of *AAT* genes and their applications in the improvement of foxtail millet.

Methods

Genome-wide scan and identification of *AAT* genes in foxtail millet

The latest foxtail millet genomic and protein sequences (*Setaria italica* v2.0, release-41) were obtained from the ensemble plants database (<http://plants.ensembl.org/info/website/ftp/index.html>). First, the *AAT* proteins identified in *Arabidopsis*, rice and wheat were aligned with the millet protein sequences by BLASTP program with an e-value of $1e^{-5}$ and an identity of 50% as the threshold. Second, the conserved domains of putative *AAT* proteins were checked using the Hidden Markov Model (HMM) profiles of the *AAT* domain (PF01490 and PF00324) from the Pfam database (<http://pfam.xfam.org/>), by a HMMER tool with a 'hmmsearch' command [62]. Third, the existence and integrity of the conserved domains in candidate proteins were further analyzed through conserved domain database (CDD) (<https://www.ncbi.nlm.nih.gov/cdd>) and InterPro database (<http://www.ebi.ac.uk/interpro/scan.html>) with default parameters. Finally, the *AAT* gene family members in foxtail millet were identified after removing those containing incomplete conservative domains and redundant sequences.

Based on the chromosome position and phylogenetic relationship, these putative *AAT* genes in foxtail millet were systematically named, and their detailed information was

counted, including cDNA length, coding sequence length, gene structure and protein sequence feature. Their gene structures and biochemical parameters were determined by GSDS (<http://gsds.gao-lab.org/>) with the genome annotation GFF3 file and Computer pI/Mw tool (https://web.expasy.org/compute_pi/), respectively [63]. As complete membrane proteins, their transmembrane (TM) regions and subcellular localization were predicted by TMHMM server 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) and LocTree 3 (<https://www.rostlab.org/services/loctree3/>) using the default parameters, respectively [64].

Chromosomal mapping, duplication and selective pressure analysis

The locations of *SiAAT* genes on the chromosomes were obtained from the ensemble plants database. The gene duplication was determined by MCScanX, and manual screening was performed according to Wang et al. [65, 66]. The distributions of *AAT* genes and their duplicated paralogous on chromosomes were plotted with TBtools software [67]. The ratios (Ka/Ks) of non-synonymous (Ka) and synonymous (Ks) of these paralogous *AAT* genes were analyzed using KaKs_Calculator 2.0 and displayed by an R package ggplot2 followed the standard processes as detailed in the software manuals [68].

Phylogenetic and conserved motifs analysis

Multi-sequence alignment of the *AAT* protein sequences identified in *Arabidopsis*, potato, rice and foxtail millet was conducted with ClustalW, and phylogenetic tree was created using the maximum likelihood (ML) method, JTT model and a bootstrap of 1000 using MEGA 6.06 [69]. The phylogenetic tree was displayed with iTOL v3 (<http://itol.embl.de/#>).

The conserved motifs were determined using Motif-based sequence analysis (MEME) program (<http://meme-suite.org>) with the parameters of motif width of 6–200, the maximum motif number of 20 and any number of repeating motif sites [70]. TM regions and the possible protein secondary structures obtained from Protein Data Bank (PDB) database (<https://www.rcsb.org/>) were annotated and displayed by ESPript 3.0 [71].

Identification of the orthologous *AAT* genes in sorghum, wheat, rice and *Arabidopsis*, and their expression analysis

The orthologous *AAT* genes in sorghum, wheat, rice and *Arabidopsis* were identified as those in foxtail millet, and by OrthFinder2 with default parameters [72]. Their expression data in the main organs of root, stem, leaf, inflorescence/spike, and grain was obtained from the public transcriptome databases, respectively (wheat, <http://www.wheat-expression.com>; rice, <http://expression.ic4r.org>; sorghum, <http://structuralbiology.cau.edu.cn/sorghum/>; *Arabidopsis*,

arabidopsis.org). Then, the conservation of the expression patterns of *AAT* orthologous genes among those species was investigated through the tissue-to-tissue and gene-to-gene correlation analysis, using the software of SPSS 19.0 (SPSS, Inc., Chicago, IL, USA) with Pearson approach.

To study the spatiotemporal expression patterns of *SiAATs* in various tissues and grain development, a merged transcriptome dataset of foxtail millet was constructed with the RNA-seq reads downloaded from Multi-omics Database for *Setaria italica* (MDSi, <http://foxtail-millet.biocloud.net/home>), including 13 tissues at the seedling, heading and filling stages, and 2 spikelets and 5 grains. The FPKM (Fragments Per Kilobase per Million) values were calculated using an R package Edge R [73]. The heat maps of \log_2 (FPKM+1) values were drawn by TBtools [67].

Quality traits measurement and transcriptome analysis of two foxtail millet cultivars

Two foxtail millet cultivars “JG21” and “YG1” used in this experiment were widely grown in northern China and provided by Mr. Junjie Wang of Center for agricultural genetic resources research, Shanxi Agricultural University. They were grown in the experimental farm of the Institute of Water Saving Agriculture in Arid Areas of China, Northwest A&F University, Yangling, Shaanxi, China (34°7′N, 108°4′E) for two cropping seasons (May to September of 2017–2018 and 2018–2019), followed the local management practices. After harvest, the mature seeds were taken for measuring quality traits, including glutenin content, crude fiber content, crude fat content, and the contents of 18 essential amino acids of methionine, cysteine, phenylalanine, alanine, glycine, glutamic acid, arginine, lysine, tyrosine, leucine, proline, tryptophan, Serine, threonine, aspartic acid, valine, isoleucine, histidine, with a near-infrared analyzer DA7250 (Perten, Sweden).

The developing grains (include glumes) of “JG21” and “YG1” were separated at filling stage in 2019, and the total RNA was isolated using a Quick RNA Isolation kit (Takara Corporation, Dalian, China) [74], RNA purity and integrity was confirmed by an RNA Nano 6000 Assay Kit and the Bioanalyzer 2100 system (Agilent Technologies, CA, USA). The RNA-seq library construction and sequencing were performed with the 150 paired-end sequencing on Illumina HiSeq platform following the standard methods by Novogene Life Sciences Pvt. Ltd., Beijing, China with three biological replicates. After filtered the raw data, Clean reads were then mapped to the reference genome using Hisat2 [75]. Differentially expressed genes (DEGs) between “JG21” and “YG1” were identified using the R package DESeq [76], and the FPKM values were calculated [77]. The KEGG

(Kyoto Encyclopedia of Genes and Genomes, <http://www.genome.jp/kegg>) and GO (Gene Ontology, <http://geneontology.org/>) annotation of DEGs were further performed with GSeq and KOBAS software, respectively [78]. All the analysis were performed following the standard methods.

The expressions of some *AAT* genes in response to abiotic stresses

To verify the reliability of the transcriptome data from RNA-seq, the flag leaves, peduncles, roots, developing grains (include glumes) of “JG21” and developing grains (include glumes) of “YG1” were collected at the filling stage in 2019 for RNA extraction. For abiotic stresses, the raw read data of two datasets (SRA062640 and PRJNA545871) of foxtail millet under drought stress and salt stress were sourced from the NCBI SRA database to analyze the response of *SiAATs* to abiotic stresses [79, 80]. The detailed analysis process was described above.

The seedlings, 15 days after sowing (two leaves with a terminal bud) were exposed to 20% polyethylene glycol (PEG 6000) solution (Guangdong Guanghua Sci-Tech Co., Ltd., Guangdong, China) and NaCl (Xilong Scientific Co., Ltd., Guangdong, China) solution (200 mM) for 1 h and 5 h to mimic salt and drought stress, respectively, and all seedling plants were then collected for RNA isolation. PrimeScript™ II 1st Strand cDNA Synthesis (Tiangen Biotech (Beijing) Co., Ltd., Beijing, China) was used to synthesize the first strand of cDNA and Quantitative Real-time PCR (qRT-PCR) was conducted by SuperReal PreMix Color kit (SYBR Green) with the primers listed in Table 3 and Additional file 7: Table S3, with the *SiAct-7* gene as an internal reference. Each sample contained three biological replicates, and the $2^{-\Delta\Delta CT}$ method was used for the statistical analysis.

Abbreviations

AAT: Amino acid transporter; APC: Amino acid-polyamine-choline; AAAP: Amino acid/auxin permease; ACT: Amino acid/choline transporter; CAT: Cationic amino acid transporter; PHS: Polyamine H⁺-symporter; TTP: Tyrosine-specific transporter; LHT: Lysine and histidine transporter; GAT: γ -aminobutyric acid transporter; ProT: Proline transporter; AAP: Amino acid permease; AUX: Auxin transporter; ANT: Aromatic and neutral amino acid transporter; GABA: γ -aminobutyric acid; QTL: Quantitative trait locus; pI: Isoelectric point; Mw: Molecular weight; TM: Transmembrane; qRT-PCR: Quantitative Real-Time PCR; GPC: Grain protein content

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-021-07779-9>.

Additional file 1: Figure S1. Physical mapping and gene duplication of *SiAAT* genes in foxtail millet. Segmental duplication and tandem duplication are linked by red lines and blue shadows, respectively.

Additional file 2: Figure S2. The sequences of the 20 conserved motifs of *SiAAT* proteins. Different colors represent different amino acid residues. The larger the font is, the higher the proportion of this residue is as in the multiple sequence alignment.

Additional file 3: Figure S3. Gene structure of 25 paralogous AAT gene groups in foxtail millet. The gene structures of the tandemly and segmentally duplicated gene groups were on the left and the right respectively. Detail legend is as in Fig. 3.

Additional file 4: Figure S4. The expression levels of the 20 selected *SiAAT* genes in five main tissues at grain filling stage. The tissue names were listed in the X axis. The mean values of three replicates \pm standard deviation (SD) were represented with Bars, and the *SiAct-7* gene is used as an internal reference.

Additional file 5: Table S1. The detail information and sequence characterization of 94 putative AAT genes in foxtail millet.

Additional file 6: Table S2. The orthologous AAT genes in foxtail millet, wheat, sorghum and *Arabidopsis*.

Additional file 7: Table S3. The primers used for the spatiotemporal expression analysis with Quantitative Real-Time PCR (qRT-PCR).

Acknowledgements

Special thanks to Mr. Junjie Wang (Center for agricultural genetic resources research, Shanxi agricultural university) for providing the foxtail millet materials used for experiments.

Authors' contributions

YY and YC performed the experiment and wrote the paper; JL collected the previous studies; JZ designed primers and performed the qRT-PCR; ZZ and CC performed transcriptome sequencing and data analysis; QL carried out quality trait investigation; AA checked the language specification of manuscript; YGH and LC designed the experiment and reviewed the paper. All authors read the final version of the manuscript.

Funding

This research was funded by the National Natural Science Foundation of China (31671695 and 31501307) and the China 111 Project of the Ministry of Education of China (B12007).

Availability of data and materials

The RNA-seq raw data of two foxtail millet cultivars "YG1" and "JG21" was submitted in the NCBI Sequence Read Archive (SRA) with the accession number PRJNA673778. The datasets supporting the conclusions of this article are included within the article and its additional files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹State Key Laboratory of Crop Stress Biology for Arid Areas, College of Agronomy, Northwest A&F University, Yangling, Shaanxi, China. ²Institute of Water Saving Agriculture in Arid Regions of China, Northwest A&F University, Yangling, Shaanxi, China.

Received: 22 November 2020 Accepted: 7 June 2021

Published online: 08 July 2021

References

- Rentsch D, Schmidt S, Tegeder M. Transporters for uptake and allocation of organic nitrogen compounds in plants. *FEBS Lett.* 2007;581(12):2281–9. <https://doi.org/10.1016/j.febslet.2007.04.013>.
- Williams LE, Miller A. Transporters responsible for the uptake and partitioning of nitrogenous solutes. *Annu Rev Plant Physiol Plant Mol Biol.* 2001;52(1):659–88. <https://doi.org/10.1146/annurev.arplant.52.1.659>.
- Williams LE, Bick JA, Neelam A, Weston KN, Hall JL. Biochemical and molecular characterization of sucrose and amino acid carriers in *Ricinus communis*. *J Exp Bot.* 1996;47(Special Issue):1211–6. https://doi.org/10.1093/jxb/47.Special_Issue.1211.
- Miranda M, Borisjuk L, Tewes A, Heim U, Sauer N, Wobus U, et al. Amino acid permeases in developing seeds of *Vicia faba* L.: expression precedes storage protein synthesis and is regulated by amino acid supply. *Plant J.* 2001;28(1):61–71. <https://doi.org/10.1046/j.1365-313X.2001.01129.x>.
- Ortiz-Lopez A, Chang H, Bush DR. Amino acid transporters in plants. *Biochim Biophys Acta.* 2000;1465(1–2):275–80. [https://doi.org/10.1016/S0005-2736\(00\)00144-9](https://doi.org/10.1016/S0005-2736(00)00144-9).
- Rolletschek H, Hosein F, Miranda M, Heim U, Gotz KP, Schlereth A, et al. Ectopic expression of an amino acid transporter (VfAAP1) in seeds of *Vicia narbonensis* and pea increases storage proteins. *Plant Physiol.* 2005;137(4):1236–49. <https://doi.org/10.1104/pp.104.056523>.
- Peng B, Kong HL, Li YB, Wang LQ, Zhong M, Sun L, et al. OsAAP6 functions as an important regulator of grain protein content and nutritional quality in rice. *Nat Commun.* 2014;5(1):4847. <https://doi.org/10.1038/ncomms5847>.
- Hunt E, Gattolin S, Newbury HJ, Bale JS, Tseng HM, Barrett DA, et al. A mutation in amino acid permease AAP6 reduces the amino acid content of the *Arabidopsis* sieve elements but leaves aphid herbivores unaffected. *J Exp Bot.* 2012;61(1):55–64. <https://doi.org/10.1093/jxb/erp274>.
- Singh BK. Plant amino acids: biochemistry and biotechnology. New York: Marcel Dekker; 1999. p. 227–47.
- Ma H, Cao X, Shi S, Li S, Gao J, Ma Y, et al. Genome-wide survey and expression analysis of the amino acid transporter superfamily in potato (*Solanum tuberosum* L.). *Plant Physiol Biochem.* 2016;107:164–77. <https://doi.org/10.1016/j.plaphy.2016.06.007>.
- Cheng L, Yuan HY, Ren R, Zhao SQ, Han YP, Zhou QY, et al. Genome-wide identification, classification, and expression analysis of amino acid transporter gene family in *Glycine max*. *Front Plant Sci.* 2016;7:515. <https://doi.org/10.3389/fpls.2016.00515>.
- Zhao H, Ma H, Yu L, Wang X, Zhao J. Genome-wide survey and expression analysis of amino acid transporter gene family in rice (*Oryza sativa* L.). *PLoS One.* 2012;7(11):e49210. <https://doi.org/10.1371/journal.pone.0049210>.
- Tian R, Yang Y, Chen M. Genome-wide survey of the amino acid transporter gene family in wheat (*Triticum aestivum* L.): identification, expression analysis and response to abiotic stress. *Int J Biol Macromol.* 2020;162:1372–87. <https://doi.org/10.1016/j.ijbiomac.2020.07.302>.
- Sanders A, Collier R, Trethewey A, Gould G, Sieker R, Tegeder M. AAP1 regulates import of amino acids into developing *Arabidopsis* embryos. *Plant J.* 2009;59(4):540–52. <https://doi.org/10.1111/j.1365-313X.2009.03890.x>.
- Okumoto S, Pilot G. Amino acid export in plants: a missing link in nitrogen cycling. *Mol Plant.* 2011;4(3):453–63. <https://doi.org/10.1093/mp/srr003>.
- Tegeder M. Transporters for amino acids in plant cells: some functions and many unknowns. *Curr Opin Plant Biol.* 2012;15(3):315–21. <https://doi.org/10.1016/j.pbi.2012.02.001>.
- Su YH, Frommer WB, Ludwig U. Molecular and functional characterization of a family of amino acid transporters from *Arabidopsis*. *Plant Physiol.* 2004;136(2):3104–13. <https://doi.org/10.1104/pp.104.045278>.
- Okumoto S, Schmidt R, Tegeder M, Fischer WN, Rentsch D, Frommer WB, et al. High affinity amino acid transporters specifically expressed in xylem parenchyma and developing seeds of *Arabidopsis*. *J Biol Chem.* 2002;277(47):45338–46. <https://doi.org/10.1074/jbc.M207730200>.
- Lee YH, Foster J, Chen J, Voll LM, Weber APM, Tegeder M. AAP1 transports uncharged amino acids into roots of *Arabidopsis*. *Plant J.* 2007;50(2):305–19. <https://doi.org/10.1111/j.1365-313X.2007.03045.x>.
- Zhang L, Tan Q, Lee R, Trethewey A, Lee YH, Tegeder M. Altered xylem-phloem transfer of amino acids affects metabolism and leads to increased seed yield and oil content in *Arabidopsis*. *Plant Cell.* 2010;22(11):3603–20. <https://doi.org/10.1105/tpc.110.073833>.
- Okumoto S, Koch W, Tegeder M, Fischer WN, Biehl A, Leister D, et al. Root phloem-specific expression of the plasma membrane amino acid proton co-transporter AAP3. *J Exp Bot.* 2004;55(406):2155–68. <https://doi.org/10.1093/jxb/erh233>.
- Boorer KJ, Fischer WN. Specificity and stoichiometry of the *Arabidopsis* H⁺ amino acid transporter AAP5. *J Biol Chem.* 1997;272(20):13040–6. <https://doi.org/10.1074/jbc.272.20.13040>.
- Schmidt R, Stransky H, Koch W. The amino acid permease AAP8 is important for early seed development in *Arabidopsis thaliana*. *Planta.* 2007;226(4):805–13. <https://doi.org/10.1007/s00425-007-0527-x>.
- Koch W, Kwart M, Laubner M, Heineke D, Stransky H, Frommer WB, et al. Reduced amino acid content in transgenic potato tubers due to antisense

- inhibition of the leaf H⁺/amino acid symporter StAAP1. *Plant J.* 2003;33(2): 211–20. <https://doi.org/10.1046/j.1365-313X.2003.01618.x>.
25. Kohl S, Hollmann J, Blattner FR, Radchuk V, Andersch F, Steuernagel B, et al. A putative role for amino acid permeases in sink-source communication of barley tissues uncovered by RNA-seq. *BMC Plant Biol.* 2012;12(1):154. <https://doi.org/10.1186/1471-2229-12-154>.
 26. Ji Y, Huang W, Wu B, Fang Z, Wang X. The amino acid transporter OsAAP1 mediates growth and grain yield by regulating neutral amino acids uptake and reallocation in *Oryza sativa*. *J Exp Bot.* 2020;71(16):4763–77. <https://doi.org/10.1093/jxb/eraa256>.
 27. Yang Y, Chai Y, Zhang X, Lu S, Zhao Z, Wei D, et al. Multi-locus GWAS of quality traits in bread wheat: mining more candidate genes and possible regulatory network. *Front Plant Sci.* 2020;11:1091. <https://doi.org/10.3389/fpls.2020.01091>.
 28. Chen L, Bush DR. LHT1, lysine- and histidine-specific amino acid transporter in *Arabidopsis*. *Plant Physiol.* 1997;115(3):1127–34. <https://doi.org/10.1104/pp.115.3.1127>.
 29. Hirner A, Ladwig F, Stransky H, Okumoto S, Keinath M, Harms A, et al. *Arabidopsis* LHT1 is a high-affinity transporter for cellular amino acid uptake in both root epidermis and leaf mesophyll. *Plant Cell.* 2006;18(8):1931–46. <https://doi.org/10.2307/20076748>.
 30. Forsum O, Svennerstam H, Ganeteg U, Näsholm T. Capacities and constraints of amino acid utilization in *Arabidopsis*. *New Phytol.* 2008;179(4): 1058–69. <https://doi.org/10.1111/j.1469-8137.2008.02546.x>.
 31. Chen L, Ortiz-Lopez A, Jung A, Bush DR. ANT1, an aromatic and neutral amino acid transporter in *Arabidopsis*. *Plant Physiol.* 2001;125(4):1813–20. <https://doi.org/10.1104/pp.125.4.1813>.
 32. Marchant A, Kargul J, May ST, Müller P, Delbarre A, Perrot-Rechenmann C, et al. AUX1 regulates root gravitropism in *Arabidopsis* by facilitating auxin uptake within root apical tissues. *EMBO J.* 1999;18(8):2066–73. <https://doi.org/10.1093/emboj/18.8.2066>.
 33. Marchant A, Bhalerao R, Casimiro I, Eklof J, Casero PJ, Bennett M, et al. AUX1 promotes lateral root formation by facilitating indole-3-acetic acid distribution between sink and source tissues in the *Arabidopsis* seedling. *Plant Cell.* 2002;14(3):589–97. <https://doi.org/10.1105/tpc.010354>.
 34. Swarup R, Kargul J, Marchant A, Zadik D, Rahman A, Mills R, et al. Structure function analysis of the presumptive *Arabidopsis* auxin permease AUX1. *Plant Cell.* 2004;16(11):3069–83. <https://doi.org/10.2307/3872286>.
 35. Swarup R, Benková E, Swarup R, Casimiro I, Péret B, Yang Y, et al. The auxin influx carrier LAX3 promotes lateral root emergence. *Nat Cell Biol.* 2008; 10(8):946–54. <https://doi.org/10.1038/ncb1754>.
 36. Grallath S, Weimar T, Meyer A, Gumy C, Suter-Grottemeyer M, Neuhaus JM, et al. The AtProT family. Compatible solute transporters with similar substrate specificity but differential expression patterns. *Plant Physiol.* 2005; 137(1):117–26. <https://doi.org/10.1104/pp.104.055079>.
 37. Lehmann S, Gumy C, Blatter E, Boeffel S, Fricke W, Rentsch D. In planta function of compatible solute transporters of the AtProT family. *J Exp Bot.* 2011;62(2):787–96. <https://doi.org/10.1093/jxb/erq320>.
 38. Mo LY, Wu LH, Tao QN. Advances in study of plant organic nitrogen nutrition. *Acta Ecol Sin.* 2002;22(1):118–24 (In Chinese).
 39. Michaeli S, Fait A, Lagor K, Nunes-Nesi A, Grillich N, Yellin A, et al. A mitochondrial GABA permease connects the GABA shunt and the TCA cycle, and is essential for normal carbon metabolism. *Plant J.* 2011;67(3): 485–98. <https://doi.org/10.1111/j.1365-313X.2011.04612.x>.
 40. Serrano R. Salt tolerance in plants and microorganisms: toxicity targets and defense responses. *Int Rev Cytol.* 1996;165:1–52. [https://doi.org/10.1016/S0074-7696\(08\)62219-6](https://doi.org/10.1016/S0074-7696(08)62219-6).
 41. Lehmann S, Funck D, Szabados L, Rentsch D. Proline metabolism and transport in plant development. *Amino Acids.* 2010;39(4):949–62. <https://doi.org/10.1007/s00726-010-0525-3>.
 42. Rentsch D, Hirner B, Schmelzer E, Frommer WB. Salt stress induced proline transporters and salt stress-repressed broad specificity amino acid permeases identified by suppression of a yeast amino acid permease-targeting. *Plant Cell.* 1996;8(8):1437–46. <https://doi.org/10.1105/tpc.8.8.1437>.
 43. Guo N, Dong X, Zhang W, Zhao J, Xue C, Yan Q, et al. Overexpression of GmProT1 and GmProT2 increases tolerance to drought and salt stresses in transgenic *Arabidopsis*. *J Integr Agric.* 2016;15(8):1727–43. [https://doi.org/10.1016/S2095-3119\(15\)61288-6](https://doi.org/10.1016/S2095-3119(15)61288-6).
 44. Breitzkreuz KE, Shelp BJ, Fischer WN, Schwacke R, Rentsch D. Identification and characterization of GABA, proline and quaternary ammonium compound transporters from *Arabidopsis thaliana*. *FEBS Lett.* 1999;450(3): 280–4. [https://doi.org/10.1016/S0014-5793\(99\)00516-5](https://doi.org/10.1016/S0014-5793(99)00516-5).
 45. He L, Zhang B, Wang X, Li H, Han Y. Foxtail millet: nutritional and eating quality, and prospects for genetic improvement. *Front Agr Sci Eng.* 2015; 2(2):124–33. <https://doi.org/10.15302/j-fase-2015054>.
 46. Li WH, Gojobori T, Nei M. Pseudogenes as a paradigm of neutral evolution. *Nature.* 1981;292(5280):237–9. <https://doi.org/10.1038/292237a0>.
 47. Javelle M, Klein-Cosson C, Vernoud V, Boltz V, Maher C, Timmermans M, et al. Genome-wide characterization of the HD-ZIP IV transcription factor family in maize: preferential expression in the epidermis. *Plant Physiol.* 2011; 157(2):790–803. <https://doi.org/10.1104/pp.111.182147>.
 48. Hudson KA, Hudson ME. A classification of basic helix-loop-helix transcription factors of soybean. *Int J Genomics.* 2015;2015:1–10. <https://doi.org/10.1155/2015/603182>.
 49. Tegeder M, Ward JM. Molecular evolution of plant AAP and LHT amino acid transporters. *Front Plant Sci.* 2012;3:21. <https://doi.org/10.3389/fpls.2012.00021>.
 50. Kong H, Landherr LL, Frohlich MW, Leebens-Mack J, Ma H, dePamphilis CW. Patterns of gene duplication in the plant SKP1 gene family in angiosperms: evidence for multiple mechanisms of rapid gene birth. *Plant J.* 2007;50(5): 873–85. <https://doi.org/10.1111/j.1365-313X.2007.03097.x>.
 51. Xiong LZ, Xu CG, Saghai Maroof MA, Zhang Q. Patterns of cytosine methylation in an elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification polymorphism technique. *Mol Gen Genet.* 1999;261(3):439–46. <https://doi.org/10.1007/s004380050986>.
 52. Downen RH, Pelizzola M, Schmitza RJ, Listera R, Downen JM, Nerya JR, et al. Widespread dynamic DNA methylation in response to biotic stress. *P Natl Acad Sci USA.* 2012;109(32):2183–91. <https://doi.org/10.1073/pnas.1209329109>.
 53. Takuno S, Gaut BS. Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. *P Natl Acad Sci USA.* 2013; 110(5):1791–802. <https://doi.org/10.1073/pnas.1215380110>.
 54. Grewal SIS, Moazed D. Heterochromatin and epigenetic control of gene expression. *Science.* 2003;301(5634):798–802. <https://doi.org/10.1126/science.1086887>.
 55. Ludwig M. Evolution of the C4 photosynthetic pathway: events at the cellular and molecular levels. *Photosynth Res.* 2013;117(1–3):147–61. <https://doi.org/10.1007/s11120-013-9853-y>.
 56. Zhang G, Liu X, Quan Z, Cheng S, Xu X, Pan S, et al. Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. *Nat Biotechnol.* 2012;30(6):549–54. <https://doi.org/10.1038/nbt.2195>.
 57. Schilling S, Kennedy A, Pan S, Jermini LS, Melzer R. Genome-wide analysis of MIK-type MADS-box genes in wheat: pervasive duplications, functional conservation and putative neofunctionalization. *New Phytol.* 2019;225(1): 511–29. <https://doi.org/10.1111/nph.16122>.
 58. Hirner B, Fischer WN, Rentsch D, Kwart M, Frommer WB. Developmental control of H⁺/amino acid permease gene expression during seed development of *Arabidopsis*. *Plant J.* 1998;14(5):535–44. <https://doi.org/10.1046/j.1365-313X.1998.00151.x>.
 59. Svennerstam H, Ganeteg U, Bellini C, Näsholm T. Comprehensive screening of *Arabidopsis* mutants suggests the lysine histidine transporter 1 to be involved in plant uptake of amino acids. *Plant Physiol.* 2007;143(4):1853–60. <https://doi.org/10.1104/pp.106.092205>.
 60. Popova OV, Dietz KJ, Gollack D. Salt-dependent expression of a nitrate transporter and two amino acid transporter genes in *Mesembryanthemum crystallinum*. *Plant Mol Biol.* 2003;52(3):569–78. <https://doi.org/10.1023/A:1024802101057>.
 61. Ueda A, Shi W, Sanmiya K, Shono M, Takabe T. Functional analysis of salt-induced proline transporter of barley roots. *Plant Cell Physiol.* 2001;42(11): 1282–9. <https://doi.org/10.1093/pccp/pce166>.
 62. Wheeler TJ, Eddy SR. Nhmmer: DNA homology search with profile HMMs. *Bioinformatics.* 2013;29(19):2487–9. <https://doi.org/10.1093/bioinformatics/btt403>.
 63. Gasteiger E, Hoogland C, Gattiker A, Se D, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy server. In: *The Proteomics Protocols Handbook*; 2005. p. 571–607. <https://doi.org/10.1385/1-59259-890-0-571>.
 64. Goldberg T, Hecht M, Hamp T, Karl T, Yachdav G, Ahmed N, et al. LocTree3 prediction of localization. *Nucleic Acids Res.* 2014;42(W1):W350–5. <https://doi.org/10.1093/nar/gku396>.

65. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, et al. MScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 2012;40(7):e49. <https://doi.org/10.1093/nar/gkr1293>.
66. Wang M, Yue H, Feng K, Deng P, Song W, Nie X. Genome-wide identification, phylogeny and expressional profiles of mitogen activated protein kinase kinase kinase (MAPKKK) gene family in bread wheat (*Triticum aestivum* L.). *BMC Genomics.* 2016;17(1):668. <https://doi.org/10.1186/s12864-016-2993-7>.
67. Chen C, Xia R, Chen H, He Y. TBtools, a Toolkit for Biologists integrating various HTS-data handling tools with a user-friendly interface. *bioRxiv.* 2018: 289660. <https://doi.org/10.1101/289660>.
68. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genom Proteom Bioinf.* 2010;8(1):77–80. [https://doi.org/10.1016/S1672-0229\(10\)60008-3](https://doi.org/10.1016/S1672-0229(10)60008-3).
69. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol.* 2013;30(12):2725–9. <https://doi.org/10.1093/molbev/mst197>.
70. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 2009; 37(Web Server):W202–8. <https://doi.org/10.1093/nar/gkp335>.
71. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 2014;42(W1):W320–4. <https://doi.org/10.1093/nar/gku316>.
72. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 2015;16(1):157. <https://doi.org/10.1186/s13059-015-0721-2>.
73. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26(1):139–40. <https://doi.org/10.1093/bioinformatics/btp616>.
74. Nicolaidis NC, Stoeckert CJ. A simple, efficient method for the separate isolation of RNA and DNA from the same cells. *Bio Techniques.* 1990;8(2): 154–6.
75. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc.* 2016;11(9):1650–67. <https://doi.org/10.1038/nprot.2016.095>.
76. Wang L, Feng Z, Wang X, Wang X, Zhang X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics.* 2009;26(1):136–8. <https://doi.org/10.1093/bioinformatics/btp612>.
77. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* 2010;28(5):511–5. <https://doi.org/10.1038/nbt.1621>.
78. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 2008; 36(Database):D480–4. <https://doi.org/10.1093/nar/gkm882>.
79. Qi X, Xie S, Liu Y, Yi F, Yu J. Genome-wide annotation of genes and noncoding RNAs of foxtail millet in response to simulated drought stress by deep sequencing. *Plant Mol Biol.* 2013;83(4–5):459–73. <https://doi.org/10.1007/s11103-013-0104-6>.
80. Pan J, Li Z, Dai S, Ding H, Wang Q, Li X, et al. Integrative analyses of transcriptomics and metabolomics upon seed germination of foxtail millet in response to salinity. *Sci Rep.* 2020;10(1):13660. <https://doi.org/10.1038/s41598-020-70520-1>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

