Genomic Survey of ATP-Binding Cassette (ABC) Transporters in *Sorghum bicolor* (L.) Moench

Olawuyi OJ¹ and Ologidi CG¹,²

¹Department of Botany, University of Ibadan, Genetics and Molecular Biology Unit, Ibadan, Nigeria
²Department of Biological Sciences, Niger Delta University, Wilberforce Island, Amassoma, Nigeria

Corresponding author: Olawuyi OJ, Department of Botany, University of Ibadan, Genetics and Molecular Biology Unit, Ibadan, Nigeria, Tel: +2348168223985, E-mail: olawuyiodunayo@yahoo.com


Received Date: February 15, 2018 Accepted Date: April 30, 2018 Published Date: May 02, 2018

Abstract

ATP-binding cassette (ABC) transporters are the largest and most ancient family of transmembrane proteins with representatives in all phyla from prokaryotes to humans. Systematic searches were performed using PSI-Blast program on NCBI to identify *Sorghum* ABC transporters. The identified proteins were subjected to phylogenetic and domain topology analyses and were classified and named according to the Human Genome Organisation (HUGO) system of classification and unified plant nomenclatural system. As part of the survey, the subcellular location, presence of signal peptide and physicochemical parameters were determined. The results from the genomic survey showed that 100 ABC transporters were identified and classified as ABCB, ABCC, ABCD, ABCE, ABCF, and ABCI. Seventy one proteins encode intrinsic membrane proteins containing nucleotide-binding and transmembrane domains (NBD and TMD) and 28 encode proteins without TMDs. The *Sorghum bicolor* ABC transporter family consists of 61 full-size molecules, 9 half-size molecules, 3 three-quarter-size molecules, and 26 quarter-size molecules. One of the proteins (GI-992164873) has neither TMD nor NBD. A majority of the proteins are located in the plasma membrane. Signal peptide was observed in SbABCB11, SbABCB34, SbABCC4, SbABCC21, and SbABCI3. The proteins had varying values of physicochemical parameters.

Keywords: *Sorghum bicolor*; ABC transporters; proteins; Genome

Introduction

*Sorghum bicolor* ranks fifth most important cereal crop in the world, and is the first plant of African origin whose genome had been sequenced. A large amount of the grains in the African continent, for instance 85 to 95% produced in Nigeria, are consumed as food; a smaller amount is used in brewery and as industrial raw materials [1].

ATP-binding cassette (ABC) transporters are largest members of diverse proteins constituting the most ancient families with representatives in all phyla from prokaryotes to humans [2]. They are transmembrane proteins that use the energy of adenosine triphosphate (ATP) binding and hydrolysis to perform certain biological processes including translocation of various solutes across membranes and non-transport-related processes such as DNA repair and translation of RNA. They transport wide ranging substances across extra- and intracellular plant membranes, including metabolic products, coating and supportive materials, defense molecules, and plant hormones that modulate the overall physiology, development, and response to biotic and abiotic stresses [3]. Thus arabidopsis ABC transport proteins such as ABCG12, ABCG11, ABCG13 have been shown to play roles in movement of cuticular lipids such that the corresponding mutants exhibits negative phenotypic and physiological aberrations [4-8]. In addition, ABC transporters including AtABCG26 and OsABCG15 have been reported to have important functions in depositing sporopollenin, which is a major component of the outer exine layer of plant spores and pollen [3,9-11]. AtABCG9 and AtABCG31 play important roles in pollen coat maturation [12].

The ABCG transporters NpPDR1/NpABC1 and NtPDR1 are involved in responds to pathogenic attacks by responding to levels of sclareol and jasmonate [13,14]. Similarly, the *Coptis japonica* ABCB1/MDR1 and ABCB2 are involved in plant defense by catalyzing the import of berberine, a potent antimicrobial compound [15,16]. ABC proteins also transport cytokinins, which facilitate communication between the shoot and root systems, abscisic acid that mediates tolerance to stress and inhibition of seed germination, auxin (indole-3-acetic acid, IAA and indole 3-butyric acid, IBA) that regulates growth and tropism, and strigolactone that is involved in plant-microbe interactions and control of shoot branching [17-25]. Furthermore ABC proteins mediate transport of metabolic intermediates and peroxisomes, which supply the required energy for germination of oil-hiving away seeds being the...
site of β-oxidation through which fatty acids are degraded, phytate (inositohexakisphosphate), a storage form of phosphate, heavy metals such as arsenic, cadmium, and mercury, and membrane lipid metabolites [26-33].

Proteins are classified as ABC transporters based on sequence homology to prototypes, phylogeny, organization of their ATP-binding cassette (ABC) domains, and presence of ABC (C) motif (also known as C motif or LSGGQ motif, (LIVMFY)S(G)GX3(RKA)(LIVMYA)X(LIVFM)) [34,35]. Plant genomes contain a plethora of ABC transporters with a few of them experimentally characterized [20]. Due to availability of resources for conducting genetic and molecular studies, Arabidopsis thaliana and Oryza sativa have received a great deal of attention allowing for experimental characterisation of several of their ABC transporters [36,37]. But only auxin transport protein (PGP1) has been experimentally described in Sorghum bicolor [38]. Therefore undertaking a genomic survey of Sorghum bicolor ABC transporters with the intent of producing an inventory of the proteins will provide an opportunity for experimentally characterising more ABC transport proteins from sorghum. Besides genomic survey of the crop's ABC transporters has not been carried out despite being the first plant of African origin whose genome had been sequenced [39].

Materials and Methods

Identification of ABC Transporters in Sorghum bicolor

Arabidopsis thaliana ABC protein sequences were retrieved from the Arabidopsis Information Resource (TAIR) database (http://www.arabidopsis.org/browse/genefamily/ABC_merged.jsp). Using the default search parameters of NCBI's Position-Specific Iterated Local Alignment Search Tool (PSI-BLAST) program (http://ncbi.nlm.nih.gov/blast), five position-specific iterative searches for S. bicolor putative/hypothetical ABC transporters were performed with Arabidopsis ABC transporter protein sequence as queries [40]. The PSI-BLAST protein-protein search program is an iterative profile method for identifying conserved patterns in BLAST results [41]. The significant threshold for selection of sequences was score of 400 with E value of e-120 [37].

Domain, Motif, Homology, and Phylogenetic Analyses

The selected sequences were subsequently subjected to domain and motif analysis-preservation of ABC domains and ABC (C) signature motifs-using default search parameters of PROSITE (http://prosite.expasy.org/), Conserved Domain Database at NCBI (http://ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) [34,35,42-44]. The domain features-forward or reverse orientation, number and size (full, three-quarter, half, and one-quarter) – were then noted. For homology analysis, prototypes of ABC subfamilies (human PDR/ABCB1 and AtABCB1 for ABCB subfamily; human ABCC1 and ABCC2 for ABC subfamily; yeast PXA1 and PXA2 for ABCD subfamily; yeast RLI for ABCE subfamily; and GCN1/AtABCF1 and GCN2/AtABCF2 for ABCF subfamily) were aligned against the selected S. bicolor ABC protein sequences by using default search parameters of blastp [45-49]. Furthermore, phylogenetic analyses were performed using neighbor joining and maximum likelihood methods with 1000 bootstrap replicates after conducting multiple sequence alignment of A. thaliana and S. bicolor ABC transporter sequences with ClustalW alignment tool [50]. Phylogenetic trees were constructed and viewed with MEGA7 and RAxM softwares programs from which similar results were obtained and only trees produced by neighbor joining method are presented and discussed herein [51,52].

Classification of Sorghum bicolor ABC Transporters into Subfamilies

The Sorghum bicolor ABC transporters identified were classified into one of six ABC subfamilies (ABCB, ABCC, ABCD, ABCE, ABCF, and ABCI) based on their domain features: half- and full-sized forward domain topology for ABCB subfamily; three-quarter- and full-sized forward topology for ABC subfamily; half- and full-sized forward orientation for ABCD subfamily; half-sized NBD-only domain topology for subfamilies ABCE and ABCF, and one-quarter-sized bacterial-like domain for subfamily ABCI [20,36,45]. Other conditions used for classifying the proteins were homology to prototype for each subfamily and clustering pattern on phylogenetic tree.

Orthology Analysis

Analysis for sequences orthologous to sorghum ABC transporters were performed using default settings of Smart Blast (https://blast.ncbi.nlm.nih.gov/smartblast/). The criteria for selecting orthologous sequences were percent identity of at least 70%, ‘E’ value of 0.0, and experimental function characterization.

Results and Discussion

Identification and Classification of ABC Transporters in Sorghum bicolor

Systematic BLAST searches of the Sorghum bicolor proteome with amino acid sequences of A. thaliana ABC transporters as queries identified 100 ABC transporters in S. bicolor that contained at least one ABC signature (Table S1). The criteria for selection of the sequences were presence of TMDs, NBDs and C motif (LSGGQ). Amongst the S. bicolor ABC transporters identified, 72 encode intrinsic membrane proteins and 28 encode proteins without TMDs. The Sorghum bicolor ABC transporter family consists of 61 full-size molecules, 9 half-size molecules, 3 three-quarter-size molecules, and 26 quarter-size molecules. One of the proteins (GI-
Sorghum bicolor have a transmembrane domain (TMD) while those in A. thaliana are 7 [36]. Twenty three (23) members of the ABCI proteins is higher in number compared to the members of ABCI proteins in the genomes of A. thaliana, O. sativa and Vitis vinifera, which contain 15, 10, and 6 [36,37,53]. Three members of the ABCI subfamily in S. bicolor consists of 2 three-quarter sized members and 24 full-sized members. Three (3) proteins were placed in ABCD subfamily because of their homology to the yeast Pxa1p and Pxa2p proteins which are the prototypic members of ABCD subfamily. The 3 proteins have a forward orientation and full-size domain structure (Table S1).

Members of ABCE and ABCF subfamily have a half-sized domain structure and are soluble proteins because they lack transmembrane domain [36]. Five (5) ABC transporters were classified as ABCE and ABCF because of their domain structure. One of the proteins showed homology to yeast Rli1p, which is a prototype for ABCF family. The other four did not show homology to Rli1p and were classified as ABC transporters (Table S1). Twenty six (26) S. bicolor ABC transporters were identified to having a single NBD or TMD (Table S1). Twenty three (23) of the proteins have a single NBD and 3 (SbABCI22, SbABCI1, SbABCI6) have a single TMD. One protein (SORBI_006G155700) with GI number 992164873 has neither NBD nor TMD but has a C motif that is characteristic of ABC transporters. Thus it was not placed in any of the subfamilies but declared as undefined (Table S1).

Forty (40) proteins of S. bicolor ABC transporters were classified as ABCB proteins because of their homology to the human ABCB1, which is a prototype member of ABCB proteins. S. bicolor ABCB transporters consist of 5 half-sized members (TAP/HMT), 2 three-quarter-sized members, and 33 full-sized members (MDR/PGP) (Table S1). Twenty six (26) S. bicolor ABC transporters belong to ABCB subfamily (Table S1). They showed homology to the human HsMRP1 and mouse cMOAT proteins which are prototypes for ABCB subfamily. ABCB transporters of S. bicolor consists of 2 three-quarter sized members and 24 full-sized members. Three (3) proteins were placed in ABCC subfamily because of their homology to the yeast Pxa1p and Pxa2p proteins which are the prototypic members of ABCC subfamily. The 3 proteins have a forward orientation and full-size domain structure (Table S1).

Plant genomes code for over 100 ABC transporters [20]. Thus the 100 ABC transporters identified in Sorghum bicolor is in accordance with the plethora of ABC transporters that have been reported from other plants including Arabidopsis thaliana, Oryza sativa, and Vitis vinifera [36,37,53] (Table S1). However, the total number of the proteins identified in S. bicolor are less than those identified in A. thaliana, Oryza sativa, and Vitis vinifera which contain 130, 121, 135, respectively [36,37,53]. This variation in number of ABC transporters in plants is a pointer that there is likely no direct proportional relationship of number of ABC transporters with genome size and chromosome number. This is further corroborated by the presence of 49 ABC transporters in humans whose genome size and chromosome number are 3,234.83 Mb and 46, respectively [3,54,55]. The large number of ABC transporters observed in Sorghum bicolor and reported in other plants is rather due to their sessile nature, multicellularity and an across-the-board capacity for synthesising secondary chemicals [56].

Sorghum bicolor genome contains less number of proteins with intrinsic transmembrane domain (Table S1) than A. thaliana, O. sativa and V. vinifera which contain 104, 108 and 120, respectively [36,37,53]. However, the Sorghum bicolor genome contains more proteins lacking TMD than A. thaliana, O. sativa, and V. vinifera, which have 18, 9 and 15, respectively [36,37,53]. In addition, the Sorghum bicolor genome contains more full-sized molecules than A. thaliana, which has 43. The ABC transporters from V. vinifera and O. sativa with full size topology of 79 and 72 are higher than the proteins from Sorghum bicolor [37,53]. The half-sized molecules identified in Sorghum bicolor are far less than those identified in A. thaliana, O. sativa and Vitis vinifera, which have 50, 41 and 41, severally [36,37,53]. However, the number of quarter-sized molecule in Sorghum bicolor is more than the numbers found in A. thaliana, O. sativa and Vitis vinifera, whose genome contain 17, 10 and 6, severally [36,37,53].

Unlike ABC transporters from A. thaliana, O. sativa and Vitis vinifera which were classified into 8 subfamilies of ABCA, ABCB, ABCC, ABCD, ABCE, ABCF, ABGG, and ABCI, the sorghum ABC transporters were classified into 6 subfamilies of ABCB, ABCC, ABCD, ABCE, ABCF, and ABCI [36,37,55]. The ABCB subfamily of sorghum contains both full-size and half-size molecules like in A. thaliana, O. sativa and Vitis vinifera. But Sorghum bicolor has 33 full-size molecules (that is, MDR/PGP) in the ABCB subfamily which is more than the members of the subfamily in A. thaliana, O. sativa and Vitis vinifera which have 21, 24, and 19, severally. The observation of 5 half size molecules of ABCB proteins in Sorghum bicolor genome shows that A. thaliana, O. sativa and Vitis vinifera contain slightly higher number of half size molecules of 7, 6, and 6, respectively.

The number of ABC (MRP) proteins identified in Sorghum bicolor is the same as those identified in Vitis vinifera with both plants having 26 members belonging to the ABC subfamily [53]. But the numbers of ABC proteins in A. thaliana and O. sativa, which are respectively 15 and 17 are lower than that in Sorghum bicolor [36,37]. The S. bicolor and O. sativa genomes have the same number of ABCD proteins with both containing 3 members [37]. The members of the ABCD subfamily is slightly lower in A. thaliana and O. sativa, which have 2 and 1, severally.

ABC and ABCC subfamily of Sorghum bicolor ABC transporters have the highest number with both containing 26 proteins. The members of S. bicolor ABCI proteins is higher in number compared to the members of ABCI proteins in the genomes of A. thaliana, O. sativa and V. vinifera which respectively contain 15, 10, and 6 [36,37,53]. Three members of the ABCI subfamily in Sorghum bicolor have a transmembrane domain (TMD) while those in A. thaliana are 7 [36]. Twenty three (23) members of the Sorghum bicolor ABCI subfamily have a nucleotide binding domain while A. thaliana genome has 10 members [36].
**Phylogeny of Sorghum bicolor ABC Transporters**

The members of the subfamilies generally clustered tightly with each other (Figure 1). Members of the ABCB subfamily clustered together tightly but two members, SbABCB15 and SbABCB14, clustered with SbABCC8, SbABCC20, SbABCC24, and SbABCC2 which are members of ABCC subfamily. In like manner, members of ABCC subfamily clustered tightly with each other but SbABCC16 and SbABCC12 formed a tight cluster with SbABCF1 while SbABCC3 grouped with members of ABCE and ABCF. Members of ABCI subfamily also clustered together tightly. However, a tight clustering of SbABCI16 with SbABCE4 and SbABCI22 with SbABCB17 were observed. The unclassified ABC transporter SORBI_006G155700 (GI: 992164873) formed a tight cluster with SbABCI8, which is a member of ABCI subfamily.

![Phylogenetic tree of S. bicolor ABC proteins.](image_url)

The generally very close clustering of members of each subfamily is as a result of the high level of similarity shared, which is due to the conserved C motif of the nucleotide binding domain (NBD) (Figure 1). The general tight clustering pattern observed in Figure 1 indicates that the nomenclatural and classification schemes adopted for Sorghum bicolor ABC transporters were effective schemes for delineating the proteins [36]. Similar clustering patterns of ABC transporters have been observed in *A. thaliana*, *O. sativa*, *Vitis vinifera* and *Lotus japonicus* [36,37,53,57].
The clustering of ShABC14 and ShABC15 with ShABCC8, ShABCC20, ShABCC24 and ShABCC2, which are members of the ABCB subfamily, is due to the similarity in the C motif sequences; LSGGQKQRIQLSRAL and LSGGQKQRIQLARAV (Figure 1). Notwithstanding, ShABC14 and ShABC15 were placed in the ABCB subfamily because of their homology to the human ABCB1, which is a prototypic member of ABCB subfamilies. In addition, similarity in C motif sequences accounts for the grouping of ShABCC16 and ShABCC12 with ShABC1 (Figure 1); LSGGQRALRLARAL, FSVGQRQLCLARAL, FSGGKMRMSLGKIL, LSGGEKRALFCFkM. The same reason can also be given for the clustering of ShABCC3 with members from other subfamilies. Nonetheless, ShABCC16, ShABCC12, and ShABCC3 were classified as ABC transporter proteins because they showed homology to the human HsMRP1 and mouse cMOAT proteins which are prototypes of ABCC subfamily.

The tight clustering of SORBI_006G155700 (GI: 992164873) with ShABC18; ShABC16 with ShABCE4; ShABC12 with ShABC17; and ShABC1 with ShABCC12 and ShABCC16 (Figure 1) can be attributed to a high similarity in their C motif sequences.

**Phylogeny of Sorghum bicolor and Arabidopsis thaliana ABC Subfamilies**

The result from phylogenetic analysis of ABCB subfamily proteins from Sorghum bicolor and Arabidopsis thaliana is shown in Figure 2. The protein sequences from A. thaliana grouped closely with those from S. bicolor with bootstrap values from 23 to 100. This is due to the high level of similarity between proteins from both plants. Eight clusters were formed between protein sequences from A. thaliana and S. bicolor. A cluster was formed around AtABCB15, AtABCB17, AtABCB16, AtABCB18, and AtABCB22. AtABCB8 (P-glycoprotein 8) formed a cluster with ShABCB20, ShABCB28, and ShABCB2. There was also clustering of AtABCB2 and AtABCB10 with ShABCB4 and ShABCB26. Another cluster observed in the phylogenetic tree was AtABCB1, AtABCB13, AtABCB14, and AtABCB19 with ShABCB8, ShABCB25 (PGP1), ShABCB6, and ShABCB5 (Figure 1). Furthermore, a cluster group was formed by AtABCB26 with ShABCB38 and ShABCB39.

Two empirically characterised ABCB proteins, AtABCB1 (alias AtPGP1) and ShABCB25 (alias ShbPGP1) which are involved in auxin efflux transport formed a cluster [38,58] (Figure 1). Thus the tight clustering of ShABCB8 with ShbPGP1 and AtPGP1 is an indication that ShABCB8 functions in auxin efflux transport. Other proteins that are likely functioning in auxin efflux transport are ShABCB6 and ShABCB9, which formed a tight clustering with AtABCB19 (alias AtPGP19, MRP11) [58]. Similarly, the function of ShABCB1 can be inferred from the function of AtABCB27 (alias aluminium tolerance-related ABC transporter, AtTAP2), which is possibly involved in redistributing internalized aluminum and may mediate vacuolar sequestration of a metal complex [36].

**Phylogeny of Sorghum bicolor and Arabidopsis thaliana ABC Subfamilies**

The result from phylogenetic analysis of ABCC subfamily proteins from Sorghum bicolor and Arabidopsis thaliana is presented in Figure 3. The protein sequences from A. thaliana grouped closely with those from S. bicolor with bootstrap values between 39 and 100. Eight clusters were formed between proteins from A. thaliana and S. bicolor. AtABCC5 and ShABCC1 formed a cluster group. AtABCC15 and AtABCC9 clustered with ShABCC22, ShABCC10, ShABCC21, and ShABCC9. AtABCC7, AtABCC6, AtABCC3 grouped with ShABCC23, ShABCC7, ShABCC26, ShABCC6, ShABCC25, and ShABCC5. Other groups comprise AtABCC11, AtABCC12, AtABCC1, AtABCC2, and ShABCC3; AtABCC13, ShABCC12, and ShABCC16; ShABCC8, ShABCC20, ShABCC2, ShABCC24, AtABCC4, and AtABCC14; AtABCC10, ShABCC14, ShABCC15, ShABCC19, ShABCC11, and ShABCC18; AtABCC8, ShABCC4, ShABCC17, and ShABCC13.

Based on clustering with ABC transporters from Arabidopsis thaliana, the function of ShABCC23, ShABCC7, ShABCC26, ShABCC6, ShABCC25, and ShABCC5 can be inferred as conjugate and heavy metal (for example, cadmium) transporters because AtABCC7 and AtABCC3 (alias MRP7) are pumps for glutathione S-conjugates and AtABCC3 (alias MRP3) transports heavy metals such as cadmium [59] (Figure 3).

In like manner, ShABCC23, ShABCC7, ShABCC26, ShABCC6, ShABCC25, ShABCC5 which formed a cluster with AtABCC7, AtABCC3, and AtABCC6 are deduced to function as pumps for glutathione S-conjugates. This is because AtABCC7 and AtABCC3 are pumps for glutathione S-conjugates [59].

The S. bicolor genome observed to having 4 members of ABCE subfamily is the plant genome with the highest number of ABCE proteins when compared with A. thaliana, O. sativa and V. vinifera whose inventory of ABC transporters have been published [36,37,53] (Table S1). However, only 1 S. bicolor ABC transporter protein was classified as ABCF protein (Table S1) as opposed to 5 in both Arabidopsis thaliana and V. vinifera and 4 in O. sativa [36,37,53].

Other Sorghum bicolor ABC proteins that are deduced to function as pumps for glutathione S-conjugate are ShABCC13, ShABCC17, and ShABCC4; ShABCC14, ShABCC15, ShABCC19, ShABCC11, and ShABCC18; ShABCC16 and ShABCC12 which formed clusters with AtABCC8, AtABCC10, and AtABCC13, respectively.

The cluster group observed between S. bicolor ABC transporter proteins, ShABCC2, ShABCC20, ShABCC8, ShABCC24 and A. thaliana proteins AtABCC14 and AtABCC4 suggests that the Sorghum bicolor proteins are functioning in pumping of glutathione.
Figure 2: Phylogenetic tree of ABCB protein sequences from *S. bicolor* and *A. thaliana*

- **S. bicolor** ABCB proteins
- **A. thaliana** ABCB proteins
S-conjugate and are involved in the regulation of stomatal aperture as a pump for folates with a high capacity. This extrapolation stems from the functions of AtABCC14 and AtABCC4 in pumping of glutathione S-conjugate and the involvement of AtABC4 (alias AtMRP4) in regulating stomatal aperture as a pump for folates with high capacity [59,60].
At ABCC1 and AtABCC2 are pumps for glutathione S-conjugates and they mediate the transport of S-conjugates including GSH, S-(2,4-dinitrophenyl)-glutathione (DNPGS), GSSG, cyaniding 3-glucoside-GS (C3G-GS) and metolachlor-GS (MOCGS), glucuronides such as 17-beta-estradiol 17- (beta-D-glucuronide) (E(2)17betaG), and the chlorophyll catabolite including *Brassica napus* nonfluorescent chlorophyll catabolite (Bn-NCC-1) [61]. Thus the formation of clusters of these *A. thaliana* ABCC transporters with SbABCC3 suggests that the *S. bicolor* ABC transporter carries out these functions.

**Phylogeny of *Sorghum bicolor* and *Arabidopsis thaliana* ABCD Subfamilies**

The result from phylogenetic analysis of ABCD subfamily proteins from *Sorghum bicolor* and *Arabidopsis thaliana* is presented in Figure 4. The ABCD transporter proteins from *A. thaliana* (AtABCD1 and AtABCD2) did not cluster closely with the proteins from *S. bicolor* with bootstrap values ranging from 0 to 100. This is might be due to marked differences between their trans membrane domain (TMD) sequences. For this reason, the *S. bicolor* ABCD transporters will most likely differ in their functions as regards the substrates that are transported.

**Figure 4:** Phylogenetic tree of ABCD proteins from *S. bicolor* and *A. thaliana* AtABCDn, *A. thaliana* ABCD proteins; SbABCDn, *S. bicolor* ABCD proteins. n: Number

**Phylogeny of *Sorghum bicolor* and *Arabidopsis thaliana* ABCE and ABCF Subfamilies**

The result for phylogenetic analysis of the soluble transport proteins (ABCE and ABCF subfamilies) from *A. thaliana* and *S. bicolor* is shown in Figure 5. Four cluster groups were formed by the protein sequences with bootstrap values of 87 to 100. One cluster group consisted of SbABCE2, AtABCF5, and AtABCF2. SbABCF1 clustered with SbABCE4, SbABCE1, and SbABCE3. The cluster formed between SbABCE2, AtABCE5, and AtABCF2 (Figure 5) cannot be used to deduce the functions of the *Soghum bicolor* ABCE transporters because the proteins from *A. thaliana* have not been empirically characterized.
Phylogeny of *Sorghum bicolor* and *Arabidopsis thaliana* ABCI Subfamilies

The result for phylogenetic analysis of ABC transporters of the ABCI subfamily from *A. thaliana* and *S. bicolor* is shown in Figure 6. Four cluster groups were formed between ABCI transport sequences from both plants. AtABCI2, AtABCI16, AtABCI15, and AtABCI10 clustered closely with SbABCI22. Similarly, AtABCI12, AtABCI13, and AtABCI17 formed a cluster group with SbABCI21, SbABCI18, SbABCI16, and SbABCI10. In addition, AtABCI13, AtABCI11, and AtABCI17 grouped together with SbABCI11. AtABCI1, AtABCI6, AtABCI8, and AtABCI9 were tightly clustered with SbABCI3.
The divergence of members of the ABC subfamily is too large making it difficult for their phylogenetic relationships to be resolved [36]. Thus tight clustering of members of the subfamily from S. bicolor and A. thaliana was not observed (Figure 6). Nevertheless AtABC13 seemed to cluster with ShABC21, ShABC18, ShABC16 and ShABC10; therefore they may have involvements in the export of heme to the mitochondrion for the biogenesis of c-type cytochromes [36]. Thus the S. bicolor ABC proteins could, with caution, be inferred to be involved in this function. In the same vein, ShABC13 could cautiously be inferred to be playing a certain part in the biogenesis of c-type cytochrome because it clustered with AtABC11 (alias cytochrome c biogenesis ABC export protein ccmA-like).

Analysis of Sequences Orthologous to Sorghum bicolor ABC Transporters

The identification of orthologs of Sorghum bicolor ABC transporters in maize and thale cress provides genetic and molecular resources for the elucidation of the biological and physiological functions of ABC transporters in sorghum. This ultimately gives an opportunity for development of better cultivars of the crop. Several orthologs to S. bicolor ABC transporters were detected but only those whose functions have been experimentally determined were used for inferring the functions of S. bicolor ABC transporters (Table S2). But it is pertinent to note that members of ABC transporter superfamily have in common the functions of active transport of substrates across membranes and hydrolysis of adenosine triphosphate (ATP) [34]. In other words they carry out ATPase activity, coupled to transmembrane movement of substances.

ShABC5 and ShABC8 are inferred to carry out auxin efflux transmembrane transporter activities and ShABCC1 is likely to function in sulfonylurea receptor activity (Table S2). ShABCC1 is also predicted to show response to stress and cellular potassium ion homeostasis (Table S2; put citation here). In addition, ShABC8, ShABC26, and SORBI_006G155700 will most likely function in fatty acid transport activities (Table S2). One member of the ABC subfamily, ShABCB25 (P-glycoprotein 1), has been characterized; it functions in calmodulin binding and auxin efflux transmembrane transporter activity.

Transport proteins planted within membranes are major targets for ameliorating the efficiency with which plants obtain and utilise water and nutrients [62]. Transport proteins are also key players in the mechanisms by which plants tolerate adverse environments including saline and acid soils. Acidic soils contain Al³⁺ ions that cause damage to the root tips of susceptible plants thereby inhibiting root growth by impaired uptake of nutrients and water [62]. Genetic studies in Sorghum bicolor have identified markers linked to Al³⁺-tolerant alleles of the S. bicolor multidrug and toxin extrusion (MATE) gene ShMATE [63]. These markers have been applied by sorghum breeders to introgress the most favourable ShMATE alleles into germplasm of sorghum. The identification of AtABCB27, which is involved in response to aluminium, as an ortholog to ShABC1 and ShABC6 adds more genetic bases to the study of Al³⁺ tolerance of sorghum and provides opportunity for increasing the genetic scope of breeding cultivars that are tolerant to Al³⁺ (Table S2). It is worthy of note to state that the result for orthology analysis of ShABC1 agrees with the result for phylogenetic analysis which showed that ShABC1 clustered tightly with AtABCB27 (Figure 2) (Table S2).

The functions of ShABC5 and ShABC8 were inferred as auxin efflux transmembrane transport based on their homology to AtABC19 and AtABC1 (alias AtPGP1) (Table S2). This agrees with the results from phylogenetic analysis which showed clustering of ShABC5 with AtABC19 and ShABC8 with AtABC1 (Figure 2). Auxin is a key player in regulation of plant growth and development; it orchestrates division, elongation, and differentiation of cells, development of embryo, tropisms in root and stem, dominance in apices, and transition to flowering [64]. Knowing that auxin transport proteins are involved in the regulation of intracellular and cell to cell auxin fluxes, the identification of ShABC5 and ShABC8 in sorghum genome could help in giving insight into the role of auxin in sorghum and monocots in general.

ShABC8 also has a promising function in calmodulin binding (Table S2). Genetic and molecular studies of the function and mechanism of action of the protein and its corresponding gene could therefore give insight into nonhost resistance, the most prevalent form of immunity in plants, in sorghum. This is because calmodulin has been implicated as a crucial requirement for nonhost resistance in A. thaliana [65].

Xenobiotics are a defense mechanism that describes the metabolism of organic compounds with abiotic origin [66]. In other words, it is a biotransformation mechanism and an essential part of plants’ strategy to cope with the potentially negative impacts of xenobiotics on growth and development. The mechanism is especially evident in the ability of plants to metabolically detoxify herbicides. There is however a differential ability of plant species to carry out this metabolic task; allowing for its extensive exploitation in modern day agriculture by use of selective herbicides that are not harmful to the crop but harmful to the associated weeds. Selective herbicidal activities come about by compounds collectively known as “herbicide safeners”, which consist of chemically many and different compounds that possess the ability to protect grass crops from injury due to herbicide application without lowering activity of herbicide in reducing target species of weeds [67]. Notwithstanding their widespread use, the molecular mechanisms employed by safeners in inducing their corresponding genes or signaling pathways are not known. Thus the identification of ShABCB9 as having a promising function in xenobiotic-transporting ATPase activity could help to uncover the molecular mechanism of induction of the corresponding gene or signaling pathways in sorghum protecting themselves from xenobiotics-based damages (Table S2). Furthermore, protection of S. bicolor from herbicide-based damages could be achieved by increasing the expression of ShABC4 genes.
The identification of maize TPA (alias lpa1) as an ortholog to SbABCC1 is an indication that the protein is involved in accumulation of phytic acid in seed grains of sorghum (Table S2). The corresponding gene of SbABCC1 could be manipulated using a site-specific technology like clustered regularly interspaced short palindromic repeats and its associated protein (CRISPR/Cas9) to down regulate or delete a segment of the gene that will cause a reduction of phytic acid in the grains. Shi et al. [27], demonstrated that targeting the gene responsible for seed/grain phytic acid accumulation in immature embryo can result in embryo-specific and eventually seed- or grain-specific phytic acid reduction. Circumventing the problem of high phytic acid content in some grains of sorghum and even in grains/seeds of other cereals and legumes by site-specific embryo-specific phytic acid reduction targeting only the gene responsible for loading of the acid into grains/seeds could be phenomenal. This is due especially to the fact that a reduction of grain/seed phytic acid that has been attempted by editing other genes in the acid biosynthetic pathway have resulted in crops with overall reduced vigour, stunted growth, reduced seed viability, and reduced yield [27,68].

Furthermore, SbABCC1 showed homology to inositol hexakisphosphate transporter of A. thaliana (AtABCC5) which agrees with the results for phylogenetic analysis with SbABCC1 forming a cluster group with AtABCC5 (Figure 3) (Table S2). AtABCC5 functions in sulfonyleurea receptor activity and shows response to salt stress via cellular potassium ion homeostasis. Thus genetic and molecular studies of the functions of SbABCC1 and its corresponding genes could help in breeding for salt tolerance in sorghum.

The results for orthology and phylogenetic analysis as regards SbABCC3 and AtABCC2 are in agreement (Table S2) (Figure 3). AtABCC2 is involved in response to cyclopentenone. Cyclopentenone take part in the biosynthesis of jasmonic acid, a defense signal employed by plants during stress [69]. The implication is that the role of cyclopentenone signals in sorghum defense against biotic and abiotic stress could be uncovered.

The functions of SbABCI8, SbABC126, and SORBI_006G155700 are inferred from their ortholog AtABCG11, which carries out functions in fatty acid transport and has involvement in cutin transport, fatty acid transport (Table S2). The homology of sorghum ABCI proteins to thale cress ABCG proteins buttresses the difficulty in resolving the phylogenetic relationships of the ABCI subfamily due to enormous diversity [36]. Cutin is a fatty substance that is deposited in cell walls and outer protective tissues of the plant body. Cutin is therefore one of the first line of defense in plants. With the identification of SbABCI8, SbABC126, and SORBI_006G155700 as fatty acid transporters in cutin formation, the mechanism and possible exploitation of cutin formation for defense could be exploited to better sorghum cultivars.

Conclusion

ABC transporters play essential role in plant organogenesis, nutrition, development, response to biotic and abiotic stress, and interaction with its environment. Thus, this study which involved an inventory of Sorghum bicolor ABC transporters could be a launch pad for unraveling the role of the proteins and corresponding genes in key biological processes in Sorghum bicolor. Therefore, there is need for further research to be conducted for elucidation of the biological and physiological functions of ABC transporters from Sorghum bicolor. Such studies could ultimately open up opportunity for improvement of Sorghum bicolor through tweaking of ABC transporter genes which are involved in organogenesis, nutrition, development, response to biotic and abiotic stress, and interaction with environment.

Supplementary Info

References

9. Shi et al. [27]. demonstrated that targeting the gene responsible for seed/grain phytic acid accumulation in immature embryo can result in embryo-specific and eventually seed- or grain-specific phytic acid reduction. Circumventing the problem of high phytic acid content in some grains of sorghum and even in grains/seeds of other cereals and legumes by site-specific embryo-specific phytic acid reduction targeting only the gene responsible for loading of the acid into grains/seeds could be phenomenal. This is due especially to the fact that a reduction of grain/seed phytic acid that has been attempted by editing other genes in the acid biosynthetic pathway have resulted in crops with overall reduced vigour, stunted growth, reduced seed viability, and reduced yield [27,68].


