RESEARCH ARTICLE

Genome-wide identification and gene expression pattern of ABC transporter gene family in Capsicum spp.

Carlos Lopez-Ortiz1●, Sudip Kumar Dutta1,2●, Purushothaman Natarajan1,3, Yadira Peña-García1, Venkata Abburi1, Thangasamy Saminathan1, Padma Nimmakayala1, Umesh K. Reddy1●*

1 Department of Biology, Gus R. Douglass Institute, West Virginia State University, Institute, West Virginia, United States of America, 2 ICAR RC NEH Region, Mizoram Centre, Kolasib, Mizoram, India, 3 Department of Genetic Engineering, School of Bioengineering, SRM Institute of Science and Technology, Kattankulathur, India

These authors contributed equally to this work.

*ureddy@wvstateu.edu

Abstract

ATP-binding cassette (ABC) transporter genes act as transporters for different molecules across biological membranes and are involved in a diverse range of biological processes. In this study, we performed a genome-wide identification and expression analysis of genes encoding ABC transporter proteins in three Capsicum species, i.e., Capsicum annuum, Capsicum baccatum and Capsicum chinense. Capsicum is a valuable horticultural crop worldwide as an important constituent of many foods while containing several medicinal compounds including capsaicin and dihydrocapsaicin. Our results identified the presence of a total of 200, 185 and 187 ABC transporter genes in C. annuum, C. baccatum and C. chinense genomes, respectively. Capsaicin and dihydrocapsaicin content were determined in green pepper fruits (16 dpa). Additionally, we conducted different bioinformatics analyses including ABC genes classification, gene chromosomal location, Cis elements, conserved motifs identification and gene ontology classification, as well as profile expression of selected genes. Based on phylogenetic analysis and domain organization, the Capsicum ABC gene family was grouped into eight subfamilies. Among them, members within the ABCG, ABCB and ABCC subfamilies were the most abundant, while ABCD and ABCE subfamilies were less abundant throughout all species. ABC members within the same subfamily showed similar motif composition. Furthermore, common cis-elements involved in the transcriptional regulation were also identified in the promoter regions of all Capsicum ABC genes. Gene expression data from RNAseq and reverse transcription-semi-quantitative PCR analysis revealed development-specific stage expression profiles in placenta tissues. It suggests that ABC transporters, specifically the ABCC and ABCG subfamilies, may be playing important roles in the transport of secondary metabolites such as capsaicin and dihydrocapsaicin to the placenta vacuoles, affecting on their content in pepper fruits. Our results provide a more comprehensive understanding of ABC transporter gene family in different Capsicum species while allowing the identification of important candidate genes related to capsaicin content for subsequent functional validation.
Introduction

Pepper (Capsicum spp.) is a member of the Solanaceae family and is closely related to potato, tomato, eggplant, tobacco and petunia. Pepper represents an important horticultural crop worldwide not only because of its economic importance, but also due to its medicinal value. Moreover, pepper fruits have been widely used as a coloring agent, food flavoring, cosmetic and pharmaceutical ingredient, and as an ornamental product. It also has nutrimental value by containing vitamins A, B, C, and E, as well as phytochemicals such as phenolic compounds, carotenoids, and capsaicin. In addition to dietary and culinary importance, capsaicinoid compounds (capsaicin and dihydrocapsaicin) of pepper have a beneficial effect for humans, including antioxidant, anticarcinogenic, antimutagenic, antiaging, and antibacterial properties [1–4].

Capsaicinoids are strongly pungent alkaloids that accumulate in the placenta of maturing Capsicum fruits. Pepper genotypes exhibit a wide range of capsaicinoid accumulation as a result of both environmental and genetic variability. Recently, Nimmakayala et al. [5] determined that the major markers linked to capsaicinoid synthesis are ankyrin-like protein, the IKI3 protein family and the ATP-binding cassette (ABC) transporter family, which suggests that their activity may be involved in the pungency modulations in pepper. The ABC transporter gene family represents one of the largest gene families, with a transporter activity that is conserved and ubiquitous in all living organisms [6, 7]. Most ABC transporter proteins that have been characterized are ATP-dependent membrane-bound transporters that are able to translocate a wide range of molecules such as lipids, proteins, chemotherapeutic drugs and heavy metals via intra- and extracellular membranes [8]. In addition, some ABC proteins also act as regulators of ion channels, receptors and proteins involved in mRNA translation and ribosome biogenesis [9].

Eukaryotes feature three common arrangements of ABC transporters: full-sized transporters composed of two transmembrane domains (TMDs) and two nucleotide-binding domains (NBDs), half-sized transporters containing one TMD and one NBD, and a third type with no TMDs but two NBDs [10]. The NBD is present in all three structural types and contains many key conserved motifs: Walker A, Q-loop, Walker B, D-loop, switch H-loop, and a signature motif (LSGGQ) that is exclusively found in ABC proteins, which distinguishes ABC proteins from other ATPases [11, 12].

Plant ABC transporters are divided into eight subfamilies—A, B, C, D, E, F, G, and I—based on their protein solubility, presence of TMDs, function, and amino acid sequence [10, 13]. Although there is a ninth subfamily group, the ABCH subfamily, it has not been identified in plants. Several ABC transporter genes have been characterized in plants, however, most of them were first identified as transporters contributing to detoxification processes [14]. Subsequently, the release of genomic data and development of bioinformatics analyses have led to comprehensive research on the identification of new functions, as well as the characterization of new ABC transporter gene family in diverse plants such as Arabidopsis, Oryza sativa and Lotus japonicus [13, 15–17], Zea mays [18, 19], Brassica rapa [20], Brassica napus [21] and Vitis vinifera [22]. Thus, apart from just detoxification functions, the understanding of several novel roles of the ABC transporter genes have been revealed. To date, it has reported that plant ABC transporters are involved in many important physiological, growth and developmental processes. Furthermore, the transport substrates of plant ABCs are divergent and include conjugated compounds, phytohormones, primary products, lipids and lipophilic compounds [23–25].

The recent release of the genome sequence of pepper [2] has provided a platform for genome-wide analysis that allow the identification and characterization of entire gene families present in hot peppers [26–28]. For instance, it makes feasible the performance of deeper studies in the ABC transporter gene family, which remains poorly understood in Capsicum spp. In
the current study, we report a genome-wide identification and characterization of ABC transporter genes in three *Capsicum* species (i.e., *C. annuum*, *C. baccatum* and *C. chinense*) including sequence alignment, phylogenetic analysis, chromosomal location and expression profile of *C. annuum* and *C. chinense*. Our results lay a foundation for further functional characterization of each ABC transporter gene among *Capsicum* species and provide useful information for better understanding the role and evolution of this gene family in higher plants.

**Materials and methods**

**Plant material**

*C. annuum* cv. CM334, *C. baccatum* cv. PBC81 and two varieties of *C. chinense* (Pimenta da neyde and Naga morich) were grown in triplicate samples in an experimental field at West Virginia State University. Fruits at 6, 16 and 25 days post-anthesis (dpa) were collected from all cultivars and stored at -80°C. Quantitative analysis of capsaicin and dihydrocapsaicin content in green pepper fruits (16 dpa) were determinate with the 1200 series HPLC system (Agilent Technologies, Santa Clara, CA) [5].

**Identification of the ABC transporter genes in pepper**

To identify all members of the ABC transporter gene family in the pepper genomes, the proteomes for the three *Capsicum* species were downloaded from the pepper genome platform (PGP) (http://passport.pepper.snu.ac.kr/?t=PGENOME) [2]. A local BLASTP search was used to query the full-length amino acid sequences of ABC transporter proteins from *Arabidopsis* (https://phytozome.jgi.doe.gov/pz/portal.html) [29]. All output genes were collected and confirmed by using the software HMMER3.0 [30]. *Capsicum* genes were searched with the PF00005 ABC transporter domain, PF01061 ABC-2 transporter domain and PF00664 ABC transporter transmembrane region domain, the ABC transporter domains were confirmed using the Pfam web server (http://Pfam.sanger.ac.uk/) [31]. Genes with E-value > 1E-05 and redundant genes were excluded. Candidate genes were analyzed in the SMART database (http://smart.embl-heidelberg.de/smart/set_mode.cgi?NORMAL=1) [32] to verify the presence of the NBD and TMD domains. Genes with NBD and TMD domains were considered members of the ABC transporter family in pepper, and the coding sequences (CDS) were downloaded from the PGP database. The Jackhammer tool (https://www.ebi.ac.uk/Tools/hmm/search/jackhammer) [33] was used to classify the ABC transporter gene family in subfamilies by using the UniProt reference proteome database with E-value = 0.01 for sequence matches and 0.03 for hit matches.

**Sequence alignment and phylogenetic analysis**

The amino acid sequences of the *Capsicum* species and *Arabidopsis* were imported into MEGAX [34] and multiple sequence alignments were carried out using ClustalW [35] with gap-open and gap-extension penalties of 10 and 0.1, respectively. The alignment file was then used to construct a phylogenetic tree based on the neighbor-joining (NJ) method. After bootstrap analysis with 1000 replicates, the tree was displayed by using the interactive Tree Of Life platform (ITOL; http://itol.embl.de/index.shtml) [36].

**Chromosomal location, Cis-element analysis and identification of conserved motifs**

The physical chromosome location data for each ABC transporter was downloaded from the PGP database and mapped onto the 12 chromosomes of pepper by using MapInspect. The
protein size, molecular weight (MW) and theoretical isoelectric point (pI) of each ABC transporter were computed by using the proteome database and sequence analysis tools on the ExPASy Proteomics Server (http://expasy.org/) [37]. For cis-element analysis, all promoter sequences (1,500 bp upstream of initiation codon “ATG”) of ABCs were extracted from the pepper genome. Then, the cis-regulatory elements of promoters for each gene were identified by using PLACE: A database of plant cis-acting regulatory DNA elements (http://www.dna.affrc.go.jp/PLACE/) [38]. Protein sequence motifs were identified by using Multiple Em for Motif Elicitation (MEME) (http://meme-suite.org/tools/meme) [39]. The analysis was performed with maximum number of motifs 10 and optimum width of motif ≥50. Discovered MEME motifs were searched in the Expasy-Prosite database with ScanProsite server (https://prosite.expasy.org/scanprosite/) [40].

Gene ontology (GO) annotation and modeling of ABC proteins
The functional annotation of ABC transporters was performed using Blast2GO software (http://www.blast2go.com). The amino acid sequences of ABC genes were imported into Blast2GO program to execute three steps: 1) BLASTp against the NCBI non-redundant protein database, 2) mapping and retrieval of GO terms associated with the BLAST results, and 3) annotation of GO terms associated with each query to relate the sequences to known protein function.

Identification of syntenic ABC paralogs pairs and gene synteny analysis
The syntenic ABC transporter paralogs pairs were identified by searching the gene duplication across all the species with the following criteria: 1) genes with >70% coverage of the alignment length; 2) genes with >70% identity in the aligned region; and 3) a minimum of two duplication events considered for strongly connected genes [41]. For each paralog pair, the non-synonymous substitution rate (Ka), the synonymous substitution rate (Ks) and the ω (= Ka/Ks) of paralog pairs were estimated by using KaKs_Calculator 2.0 [42]. The duplication date of paralog pairs was estimated by the formula T = Ks/2λ, assuming a clock-like rate (λ) of 6.96 synonymous substitutions per 10^{-9} years [43].

Transcriptome sequencing of C. chinense green fruits
Green fruits (16 dpa) from two different cultivars of C. chinense were used for whole-transcriptome sequencing. Total RNA was isolated from the pooled tissues of three biological replicates for each cultivar with the Plant RNA mini spin kit (Macherey-Nagel). The quantity and quality of the total RNA were analyzed with the Agilent 2100 Bioanalyzer and Qubit 4 Fluorometer (Invitrogen), respectively. The RNA sequencing libraries were prepared by using the NEBNext Ultra II RNA Library Prep Kit according to the manufacturer’s protocol. The mRNAs were enriched by using magnetic beads with Oligo (dT), then fragmented into shorter fragments with a fragmentation buffer. The first-strand cDNA was synthesized from the fragmented mRNA with a random hexamer primer. The resulting cDNAs were added to sequencing adapters, and sequencing primers were used for library amplification. The insert size of the library was analyzed with Agilent 2100 Bioanalyzer (Invitrogen), and the Qubit 4 Fluorometer (Invitrogen) was used for library quantification. The RNA sequencing library from each sample was sequenced in the Illumina NextSeq 500 platform with paired-end sequencing. The resulting image files were converted to FASTQ with 2x75-bp reads. The Illumina reads were deposited with the Sequence Reads Archive (NCBI) under the following accession number PRJNA526219.
Analysis of *C. chinense* transcriptome to study ABC transporter genes

The sequencing adapters and low-quality reads (Phred score QV<30) were removed by using cutadapt ([https://cutadapt.readthedocs.io/en/stable/guide.html](https://cutadapt.readthedocs.io/en/stable/guide.html)) and sickle ([https://github.com/najoshi/sickle](https://github.com/najoshi/sickle)) respectively. The quality-filtered reads were mapped to the *C. chinense* reference genome ([2]) by using the mem algorithm of the BWA tool ([46]) to generate SAM alignment. The read count table for genes from *C. chinense* was created for all the samples by using the SAM alignment and HTSeq R package ([47]). The gene expression based on the read counts were studied by reads per kilobase per million (RPKM). The RPKM values for each gene were calculated based on the read count table, the total number of reads and gene length (kb). The ABC transporters in *C. chinense* (CcABCs) were identified by homology search against the CDS sequences from *C. annuum* by using a BLASTN algorithm (identity ≥ 98% and coverage ≥ 70%). The gene annotation of the ABC transporter genes identified from *C. chinense* was confirmed by using the BLASTx algorithm against the NCBI non-redundant protein database.

Expression pattern of ABC transporters in *C. annuum* and *C. chinense*

The RNA-seq gene expression data in placenta tissues (6 dpa, 16 dpa, 25 dpa) from *C. annuum* cv. CM334 was retrieved from the RNA-seq data published by [2]. A BLASTN search was performed (identity ≥ 98% and coverage ≥ 70%) to identify the orthologs genes between *C. annuum* ABC (CaABC) and *C. chinense* (CcABC) transporters. The RPKM expression values for identified CaABC protein genes were extracted from the dataset and a gene expression heatmap was generated for *C. annuum* and *C. chinense* orthologs by using the ClustVis web tool ([https://biit.cs.ut.ee/clustvis/](https://biit.cs.ut.ee/clustvis/)) ([48]).

RNA isolation and quantitative real-time PCR (qRT-PCR)

Total RNA was isolated from pepper fruits (6, 16 and 25 dpa) by using the Plant RNA mini spin kit (Macherey-Nagel). First-strand cDNA was synthesized with 1 μg total RNA per sample by using the Super Script First-Strand Synthesis system (Invitrogen). To identify in the three *Capsicum* genomes the orthologs of the markers previously reported by [5] for the ABC transporter family, the CDS sequences for the CA06g14430 and CA11g09150 genes were downloaded from the Sol Genomics database ([https://solgenomics.net/](https://solgenomics.net/)) ([49]) and a BLASTN search was performed (identity ≥ 98% and coverage ≥ 70%) across the three pepper genomes. Gene-specific primers for the selected *Capsicum* ABC transporter orthologs were designed by using Primer3Plus ([http://www.primer3plus.com/](http://www.primer3plus.com/)). The qRT-PCR analysis involved a StepOnePlus Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) with a total volume of 20 μL containing 1 μL cDNA template, 2 μL forward and reverse primers (10 μM), 10 μL SYBR Green PCR Master (ROX) (Roche, Shanghai) and 7 μL sterile distilled water. For each sample, three replicates were run to compute the average Ct values. The data were analyzed by the 2−ΔΔCt method ([50]). Relative gene expression was normalized against that of the endogenous control β-tubulin ([51]).

Results and discussion

Capsaicin and dihydrocapsaicin content in pepper

Capsaicinoids are responsible for the hot or burning sensation of chili, pungency and flavor are the primary properties of pepper fruits ([52]). About 80% to 90% of capsaicinoids in chili fruit is represented by capsaicin and dihydrocapsaicin, and their accumulation occurs over a relatively short period during the latter stages of fruit development ([53]). *C. chinense* is one of
the hottest chili peppers in the world; in general, chili species and varieties contain about 1% capsaicin, but this content can range from 2% to 4% [54]. In this study, the highest capsaicin and dihydrocapsaicin content was for *C. chinense* cv. Naga morich, with 14.67 mg g\(^{-1}\) and 5.54 mg g\(^{-1}\) dry weight (DW) tissue, respectively. On the other hand, 4.62 mg g\(^{-1}\) and 1.08 mg g\(^{-1}\) DW tissue were reported for *C. chinense* cv. Pimienta da neyde, and 0.823 mg g\(^{-1}\) and 0.393 mg g\(^{-1}\) DW tissue in *C. annuum* cv. CM334. The lowest value across all the species was for *C. baccatum*, with a content of 0.55 and 0.15 mg g\(^{-1}\) for capsaicin and dihydrocapsaicin respectively (Fig 1).

Capsaicinoids biosynthesis is carried out principally in the placental tissues of pepper fruits by the action of several enzymes [55, 56]. Recently, NGS approaches including genotyping by sequencing (GBS), based GWAS and RNAseq analysis of placenta tissues have been used for the identification of novel genes involved in the capsaicinoids biosynthesis pathway. Moreover, these approaches have allowed the study of the mechanisms involved in the pungency modulations in pepper. Liu et al. [57] predicted the function of three novel genes i.e., dihydroxyacid dehydratase (DHAD), threonine deaminase (TD) and prephenate aminotransferase (PAT) which play key roles in the capsaicinoids biosynthetic pathway. In a recent association mapping study carried out by Nimmakayala et al. [5], it was identified significant SNPs associated with capsaicin content and fruit weight. This study revealed that genes such as Ankyrin-like protein, IKI3 family protein, pentatricopeptide repeat protein and ABC transporter G and C subfamilies are important players regulating capsaicin content. The SNPs associated with the ABC transporter gene family were S6_203416571 and S11_83592400 in the locus CA06g14430 and CA11g09150 respectively (Fig 2). Particularly, the SNP S6_203416571 located in chromosome 6, showed a high allelic effect (Fig 2A).

**Genome-wide identification of ABC proteins in pepper**

To identify the ABC protein family in pepper, we performed a BLASTP search of the three pepper genomes from the PGP database. A total of 572 genes potentially encoding ABC proteins were identified: 200 from *C. annuum* (CaABC), 185 from *C. baccatum* (CbABC), and 187 from *C. chinense* (CcABC) (Table 1). To investigate the evolutionary relationship between *Capsicum* species and *Arabidopsis* ABC transporter proteins (AtABC), we performed phylogenetic analysis of the pepper and *Arabidopsis* ABC proteins. The protein sequences of *Capsicum* ABC genes and AtABC proteins (119 protein sequences containing the ABC transporter

![Fig 1. Capsaicin and dihydrocapsaicin content in pepper.](https://doi.org/10.1371/journal.pone.0215901.g001)
domain) were aligned by using MEGAX, and an unrooted phylogenetic tree was constructed by a NJ method with 1000 bootstrap replications (Fig 3).

An extensive research on ABC transporters has resulted in several naming schemes. In most of the cases, the transporters were named on the basis of mutant characteristics. Thus, different names were assigned to the same subfamily or selected members with common characteristics. To conform to plant and animal ABC communities, the Human Genome

![Fig 2. Allelic effect of two significantly associated SNPs markers for capsaicin content in C. annuum.](https://doi.org/10.1371/journal.pone.0215901.g002)

Table 1. Comparative analysis of ABC proteins between Capsicum and other plant species.

| Species name                  | ABCA | ABCB | ABCCC | ABCD | ABCE | ABCF | ABCG | ABCI | Reference |
|-------------------------------|------|------|-------|------|------|------|------|------|-----------|
| Arabidopsis thaliana          | 12   | 29   | 15    | 2    | 3    | 5    | 43   | 21   | 130       |
| Arabidopsis lyrata            | 12   | 25   | 15    | 2    | 3    | 10   | 43   | 22   | 132       |
| Brassica rapa                 | 11   | 38   | 21    | 2    | 7    | 7    | 63   | 30   | 179       |
| Populus trichocarpa           | 6    | 40   | 29    | 3    | 2    | 4    | 78   | 42   | 204       |
| Glycine max                   | 8    | 49   | 40    | 10   | 2    | 10   | 113  | 39   | 271       |
| Carica papaya                 | 5    | 18   | 13    | 3    | 2    | 4    | 36   | 32   | 113       |
| Vitis vinifera                | 5    | 30   | 26    | 1    | 1    | 6    | 71   | 41   | 181       |
| Brachypodium distachyon       | 6    | 32   | 19    | 4    | 4    | 7    | 44   | 22   | 138       |
| Amborella trichopoda          | 4    | 19   | 17    | 2    | 2    | 5    | 46   | 42   | 137       |
| Oryza sativa                  | 6    | 28   | 17    | 3    | 1    | 6    | 56   | 24   | 141       |
| Oryza sativa                  | 6    | 27   | 17    | 3    | 2    | 6    | 50   | 16   | 127       |
| Zea mays                      | 6    | 31   | 13    | 4    | 2    | 7    | 54   | 13   | 130       |
| Brassica napus                | 30   | 69   | 47    | 5    | 13   | 14   | 116  | 20   | 314       |
| Ananas comosus                | 5    | 20   | 16    | 2    | 1    | 5    | 42   | 9    | 100       |
| Solanum lycopersicum          | 9    | 29   | 26    | 2    | 2    | 6    | 70   | 10   | 154       |
| Capsicum anuum                | 10   | 48   | 26    | 2    | 1    | 10   | 95   | 8    | 200       |
| Capsicum baccatum             | 7    | 41   | 24    | 3    | 1    | 6    | 94   | 9    | 185       |
| Capsicum chinense             | 9    | 44   | 23    | 5    | 1    | 6    | 91   | 8    | 187       |

https://doi.org/10.1371/journal.pone.0215901.t001
Organization (HUGO) nomenclature system [10] was adopted to designate all putative ABC proteins as ABCA-G and ABCI to all ABC transporter subfamilies. Overall, Capsicum ABC proteins followed the same pattern as Arabidopsis (Fig 3). Based on phylogenetic association with AtABCs and using the jackHmmer tool, Capsicum ABCs were classified into eight subfamilies previously mentioned. The number of members of ABCs within each subfamily in Capsicum were similar to other plants such as Arabidopsis [13], B. rapa [20] and tomato [59]. In order of abundance, ABCG, ABCB and ABCC subfamilies were the most prevalent groups throughout all species, whereas the smallest number of members were in the ABCD and ABCE subfamilies; for this last subfamily, only one member was identified in all the three Capsicum species analyzed.

For convenience, the ABC transporters were named CaABC1 to CaABCn for C. annum based on their subfamily group and were classified similarly for the other species. The Capsicum ABC proteins vary substantially in size and sequences of their encoded region, as well as in their physicochemical properties across all species. The locations of the ABC domains within the protein also differ. The physical locations, coding sequence length, protein characteristics and topology for ABC transporters identified for each species are in S1–S3 Tables. The domain organizations for ABC transporters are almost as varied as their function: proteins of the ABCA-ABCD subfamilies have a forward direction for domain organization (TMD-NBD), whereas the proteins of the ABCG and ABCH subfamilies contain the reverse domain organization (NBD-TMD). ABCE and ABCF proteins contain only two NBDs and were characterized as soluble proteins. ABCI proteins generally possess only one domain, mainly NBD or

Fig 3. Phylogenetic relationships of Capsicum species and Arabidopsis ABC transporter proteins. The 572 and 119 ABC proteins identified from Capsicum species and Arabidopsis, respectively, were subjected to phylogenetic analysis by the neighbor-joining method with 1000 bootstrap replicates. Subfamily names (ABCA-I, except ABCH) are indicated by different colors.

https://doi.org/10.1371/journal.pone.0215901.g003
TMD. Topological diversity is one of the unique characteristics of ABC proteins. The ABC transporters are divided in three common arrangements: full-sized transporters, half-sized transporters and a third type that has no TMDs but two NBD domains [10]. A typical full-sized ABC protein consists of ≥1,200 amino acid residues [14]. The 200 CaABC proteins ranged from 52 to 1831 amino acid residues, the CbABCs from 89 to 1864 residues and the CcABCs from 86 to 1965 residues. Nevertheless, it is important to mention that all of them possess at least one NBD, thus, they can be classified as ABC transporters and were included in this study. Some of the pepper ABC proteins with shorter sequences might be thought as pseudogenes or not annotated genes. These shorter sequences were also found in the genome-wide analysis of ABC transporters in tomato, B. rapa and pineapple [20, 58, 59]. Among the 572 ABC transporters, 212 lack a TMD and were considered soluble ABC proteins. The remaining 360 members possess TMDs and were considered ABC transporters across all species. Overall, 134 Capsicum ABC proteins are full-sized proteins possessing (TMD-NBD)x2 domains: 46, 40, 48 for C. annuum, C. baccatum and C. chinense, respectively. Among these members, 22, 24 and 28, respectively, exhibit a forward topology (TMD-NBD), whereas 24, 26, and 20 have a reverse topology (NBD-TMD). In total, 135 ABC transporters were classified as half-sized, having forward (TMD-NBD) or reverse (NBD-TMD) orientations. Among the half-sized Capsicum ABC proteins, 26 exhibit a forward and 109 a reverse domain orientation. A total of 233 ABC transporters were considered quarter-sized or single-structure proteins: 184 have an NBD domain, and 49 a TMD domain. Capsicum ABC proteins were also classified under an ABC2 (NBD-NBD) structure: 26 have the NBD-NBD structure and 3 the TMD-TMD structure. In total, 37 ABCs were uniquely characterized, with NBD-TMD-NBD, TMD-NBD-TMD and TMD-TMD-NBD-TMD-TMD structures. The differences in the topology domain orientations might have resulted from gene duplication during evolution or evolved to render specific physiological functions under biotic or abiotic stress [60].

Chromosomal locations and syntenic Capsicum ABC paralog pairs

A total of 544 (95.1%) ABC transporters were physically mapped on all 12 chromosomes of pepper, and the other 28 genes were located on unanchored scaffolds (Fig 4). ABCG, ABCB and ABCC subfamilies are unevenly distributed across all chromosomes. ABCD (in chromosome 2 and 12) and ABCE (in chromosome 01) subfamilies are the most conserved across all the Capsicum species. Among all chromosomes, chromosome 3 of C. annuum contains the highest number of ABCs—32 (16%)—followed by chromosome 6 (14.5%). Among all species, chromosomes 3, 6 and 12 contain the highest number of ABCs, with the minimum on chromosome 10.

The distribution pattern of ABC transporters on individual chromosomes also indicated certain physical regions with a relatively higher accumulation of multiple ABC gene clusters, such as chromosome 3 and 6 at the lower end of the arms for all species. The distribution of ABC transporters differs among the three genomes. Some ABC gene clusters occur in one species but not in the other genomes; for example, in chromosome 2, ABCs were present in the upper chromosome part in C. annuum and C. baccatum but were absent in C. chinense. On the other hand, ABCG and ABCB family members were found on the lower part of chromosome 4 in C. chinense and C. baccatum but not in C. annuum. A clear example is at the upper end of chromosome 8, where a cluster of genes corresponding to the ABCB and ABCG members are present in C. chinense, and only one gene appears in C. baccatum, but with no presence reported in C. annuum.

Syntenic paralogs are genes that are located in syntenic fragments. The syntenic paralog pairs were identified between and within the three Capsicum genomes. Simultaneously, we
Fig 4. Chromosomal locations of ABC transporter proteins in pepper *C. annuum* (green), *C. baccatum* (blue) and *C. chinense* (orange). Chromosome numbers are represented at the top of each chromosome. The left panel scale
further identified synonymous (Ks) and non-synonymous (Ka) values to explore the selective pressures on these paralog pairs to understand the expansion of this gene family in pepper. In total, 14 pairs of ABC transporter syntenic paralogs were identified across all species (Table 2). Six paralog pairs were within species and the remaining were intra-species. Among the eight intra-species duplications, three segmental duplication gene pairs were intra-chromosomal, located on chromosomes 5 and 3 for *C. baccatum* and chromosome 6 for *C. annuum*. Only one segmental duplication CaABCF8-CaABCF2 in *C. annuum* involved two different chromosomes. Moreover, the duplicated paralog ABC transporter pairs belong to the same subfamily. The Ka/Ks (ω) ratios for segmental duplications ranged from 0.06 to 1.57, with a mean of 0.81. In total, 11 out of 14 of the paralogs pair were under purifying selection, with ω ratios < 1. The ω ratios for 3 syntenic paralogs (21.42%) were > 1, which indicates a positive selection on these paralogs. The CaABCF8-CaABCF2 pair had the highest ω ratios with 1.57.

The duplication time of *Capsicum* ABC paralog pairs was estimated by using a relative Ks measure as a proxy for time, and it spanned from 1 to 84 million years ago (MYA), with an average duplication time of ~26 MYA. Multiple copies of genes in a gene family could have evolved due to the flexibility provided by events of whole-genome tandem and segmental duplications. Gene duplication, segmental or tandem, has been documented in several plant gene families, such as NAC, MYB, F-box, bZIP and ABC transporters [20, 61]. The ω ratios for 3 pairs of paralogs were > 1, representing positive selection and fast evolutionary rates in these ABC paralogs at the protein level. This finding differs from other gene families in plants, such as BURP in Medicago and ACD in tomato, which contain a few or even no paralog pairs undergoing positive selection [62, 63]. In our study, a relatively large percentage (~ 21%) of ABC paralogs pairs underwent positive selection. We assumed that these paralog gene pairs might have evolved in order to acquire new functions and adjust to their living environment.

### Table 2. Ka-Ks calculation of each pair of syntenic Capsicum ABC paralogs.

| Syntenic paralog pairs | S-Sites | N-Sites | Ka | Ks | Ka/Ks | Selection pressure | Duplication time (MYA) |
|------------------------|---------|---------|----|----|-------|-------------------|------------------------|
| CbABCG11-CbABCG40      | 127.23  | 406.77  | 0.23| 0.21| 1.06  | Positive selection | 15.36                  |
| CaABCG2-CaABCG40       | 142.31  | 490.69  | 0.23| 0.25| 0.89  | Purifying selection| 18.13                  |
| CaABCF8-CaABCF2        | 91.15   | 310.85  | 0.07| 0.05| 1.57  | Positive selection | 3.25                   |
| CaABCG5-CaABCG22       | 88.18   | 256.82  | 0.11| 0.15| 0.72  | Purifying selection| 11.12                  |
| CbABC3-CaABC9          | 260.28  | 843.72  | 0.00| 0.02| 0.06  | Purifying selection| 1.40                   |
| CaABC31-CaABC3         | 326.66  | 996.34  | 0.07| 0.08| 0.84  | Purifying selection| 6.00                   |
| CcABCG3-CcABCG34       | 49.75   | 166.25  | 0.10| 0.09| 1.13  | Positive selection | 6.11                   |
| CcABCG30-CaABCG80      | 108.39  | 371.61  | 0.52| 0.54| 0.95  | Purifying selection| 39.06                  |
| CbABCG89-CaABCG69      | 497.18  | 1704.82 | 0.02| 0.08| 0.31  | Purifying selection| 5.47                   |
| CaABCG23-CaABCG33      | 401.44  | 1314.56 | 0.01| 0.01| 0.47  | Purifying selection| 0.99                   |
| CaABCG7-CaABCG38       | 84.41   | 281.59  | 0.17| 0.20| 0.84  | Purifying selection| 14.57                  |
| CbABCG72-CaABCG73      | 52.13   | 184.87  | 0.82| 1.00| 0.82  | Purifying selection| 72.07                  |
| CbABCG29-CaABCG30      | 95.05   | 327.95  | 0.96| 1.16| 0.83  | Purifying selection| 83.55                  |
| CbABCG58-CaABCG38      | 148.66  | 514.34  | 1.01| 1.15| 0.88  | Purifying selection| 82.39                  |

S-Sites, number of synonymous sites; N-Sites, number of non-synonymous sites; Ka, non-synonymous substitution rate; Ks, synonymous substitution; MYA, million years ago.
Expression correlation analysis of syntenic ABC paralog pairs across different tissues and under stress treatments could help to reveal their functional roles in evolutionary fates.

**Motif composition and Cis-elements of Capsicum ABC genes**

MEME analysis according to domain composition of pepper ABC transporter proteins revealed 10 conserved motifs in ABCA-G and ABCI families (Fig 5 and S4 Table). The lengths of the conserved motifs ranged from 15 to 50 amino acids. Additionally, the number of conserved motifs in each Capsicum ABC transporters ranged from 1 to 8. The information obtained from ScanProsite analysis revealed that the function of most of the motifs was pleiotropic drug resistance related to the ABCG subfamily. All conserved motifs predicted have similar properties as ABC transporters, and the signature motif (LSGGQ) was found in most of the Capsicum ABC transporters.

In order to identify putative cis-elements in the Capsicum ABC promoters, 1500 bp DNA sequences upstream of the start codon (ATG) for the ABC transporters for each species were analyzed by using the Plant Cis-acting Regulatory DNA Elements (PLACE) website. The analysis identified 124 different cis-elements in all Capsicum ABC transporters. A total of 23 common cis-regulatory elements were present across all the promoter regions of the ABC transporters and were highly conserved among all Capsicum species (Table 3).

Four common cis-regulatory elements, CATATGMSAUR, ASF1MOTIFCAMV, NTBBF1 ARROLB and ARFAT, were found related to plant hormones including auxin, auxin response factor (ARF) and Small Auxin-Up RNAs (SAUR), which suggests that these plant hormones could affect the expression of Capsicum ABC transporters and can affect the plant growth and development. The WRKY71OS cis-regulatory element is responsive to stresses caused by pathogens. Out of the 23 common cis-regulatory elements, TBOXATGAPB, BOXIIPCCHS, INRNTPSADB, and GT1CONSENSUS are thought to be required for transcriptional
regulation by light. Two common cis-elements, CCAATBOX1 and LTRECOREATCOR15 were identified to response to low temperature, cold, drought and heat shock, which suggests that Capsicum ABC transporters might be involved in response to abiotic stress.

**GO annotation of ABC transporter genes**

GO analysis performed with Blast2Go suggested the putative participation of ABC genes in multiple biological processes, molecular functions, and cellular component (Fig 6). GO results

![Figure 6](https://doi.org/10.1371/journal.pone.0215901.g006)

Fig 6. Detailed gene ontology analysis results for Capsicum species. Biological process, cellular component, and molecular function were identified with the Blast2GO program.

https://doi.org/10.1371/journal.pone.0215901.g006

Table 3. Common putative cis-elements identified in the promoter sequences of ABC proteins genes in Capsicum species.

| Cis-element                | Signal sequence | SITE     | Expression pattern                                                                 |
|---------------------------|-----------------|----------|-------------------------------------------------------------------------------------|
| ABRELATERD1               | ACGTG           | S000414  | ABRE, etiolation, erd                                                                |
| GCN4OSGLU1                | TGAGTCA         | S000277  | GluB-1, glutelin, endosperm, seed, storage protein, GCN4 motif                      |
| TATATABOX4                | TATATAA         | S000111  | TATA, sporamin, phaseolin                                                            |
| CATATGM1SAUR              | CATATG          | S000370  | SAUR, NDE, auxin                                                                   |
| ASF1MOTIFCAMV             | TGACG           | S000024  | TGACG, root, leaf, CaMV, 35S, promoter, auxin, salicylic acid                        |
| NTBBF1ARROLB              | ACTTAA          | S000273  | rolB, Dof, auxin, domain B, root, shoot, meristem, vascular                         |
| ARFAT                     | TGCTC           | S000270  | auxin, AuxRE, ARF, ARF1, Aux/IAA, SAUR, NDE, GH3, D1, D4                            |
| CAATBOX1                  | CAAT            | S000028  | CAAT, legA, seed                                                                    |
| CCAATBOX1                 | CCAAT           | S000030  | HSE (Heat shock element), CCAAT box                                                 |
| HEXMOTIFTAH3H4            | ACGTCA          | S000053  | hexamer, HBP-1A, HBP-1B, histone H3, CaMV, 35S, NOS, HBP-1                           |
| T/GBOXATPIN2              | ACGTGT          | S000458  | T/G-box, JA, pin2, LAP, MYC, wounding                                               |
| TATCCAYM1TOSRAMY3D        | TATCCAY         | S000256  | GATA, amylase, sugar, repression                                                    |
| LTRE1HVBLT49              | CCAGAA          | S000250  | low temperature, LTRE                                                               |
| GT1CONSENSUS              | GRWAAN          | S000198  | GT-1, light, TATA, TFIIA, TBP, HR, SAR, TMY, leaf, shoot                            |
| INRNTP4SDB                | YTCANTY         | S000395  | initiator, light-responsive transcription, TATA-less promoter                       |
| TATCCAOSEARCH             | TATCCA          | S000403  | alpha-amylase, MYB proteins, gibberellin, GA, sugar starvation                     |
| TACGTVMMY                 | TGACGT          | S000377  | alpha-Amylase, cotyledon, seed germination, seed                                   |
| ABREATCONSENSUS           | YACGTGCC        | S000406  | ABA, ABF, bZIP factors                                                             |
| BOX1IPCCHS                | ACGTGCC         | S000229  | Box II, Box 2, CHS, chs, light regulation                                           |
| LRENRPCABE                | ACGTGCCA        | S000231  | CAB, cab, cab-E, CABE, light, leaf, shoot                                           |
| WRKY17OS                  | TGAC            | S000447  | WRKY, GA, MYB, W box, TGAC, PR proteins                                             |
| LTRECOREATCOR15           | CGAC            | S000153  | low temperature, cold, LTRE, drought, ABA, cor15a, BN115, leaf                     |
| TBOXATGAPB                | ACTTTG          | S000383  | GAPB, glyceraldehyde-3-phosphate dehydrogenase, light-activated                    |

https://doi.org/10.1371/journal.pone.0215901.t003
indicated the putative participation of Capsicum ABC transporters in transmembrane transport as a principal biological process, as well as drug transmembrane transport, xenobiotic transport and DNA integration. ATP binding and ATPase activity coupled to transmembrane of substances were the main activities for molecular function. Most of the ABC transporters were classified in the integral component of the membrane for cellular localization followed by the plasma membrane. In all species, 18 ABC transporters from C. chinense, 12 from C. annuum and 12 from C. baccatum were cellular localized in the vacuolar membrane and plant-type vacuole. In pepper fruit, capsaicinoids are synthesized exclusively in placental tissue and accumulate in vacuoles of placental epidermal cells [64], so ABC transporters might participate in vacuolar capsaicinoid uptake and transport, affecting the capsaicinoid content in pepper fruits.

Capsaicinoid levels are highly dynamic during fruit development. Their levels appear to be influenced by the ontogenetic trajectory of the fruit. Capsaicinoids begin to accumulate from the early stages (10 dpa) of fruit development, peak at about 40 dpa, and then it decreases sharply [65]. The late decrease in capsaicinoid content appears to result from high peroxidase activity, which oxidizes capsaicinoids in the presence of hydrogen peroxide (H₂O₂) [66, 67]. A gene CcABCC12 from C. chinense was found to have a H₂O₂ catabolic process as a biological process resulting in the breakdown of H₂O₂ (S5 Table), which suggests a detoxification process of H₂O₂ exclusively for C. chinense and a subsequent high content of capsaicinoids. Another factor that can affects the metabolism of capsaicinoids is mineral nutrition. Nitrogen (N) and potassium (K) are the main mineral players. Nitrogen availability in soil directly affects capsaicin accumulation since a single capsaicin molecule synthesis involves three amino acids such as phenylalanine, valine and leucine [68]. By contrast, potassium does not participate in capsaicinoid metabolism, however it has been reported that an increase in potassium concentration significantly decreases the capsaicin levels and leaf nitrogen content in C. chinense [69]. Thus, the level of potassium might indirectly affect capsaicin accumulation via its effects on fruit development [70]. CcABCC1 in C. chinense showed cellular potassium ion homeostasis as a biological process. The principal function of this biological process involves maintenance of an internal steady state of potassium ions at the level of a cell. In fact, C. chinense was found to have the highest values for capsaicin and dihydrocapsaicin, suggesting that cellular potassium ion homeostasis may indirectly affect capsaicinoid levels in pepper fruits.

**Expression profile of ABC transporters in C. annuum and C. chinense**

A BLASTN strategy was used to identify the orthologs for Capsaicinoid markers previously identified by [5] for the CA06g14430 gene from the SGN database. The resulted orthologs were CaABCG28, CbABCG26, and CcABCG37 corresponding to each of the species. For CA11g09150, the orthologs were CaABCC9, CBABCC5 and CcABCC20. The main purpose of gene expression profiling is to determine the genes that are differentially expressed within the organism being studied. In the same way, we used a BLASTN search to identify the orthologs between C. annuum and C. chinense to correlate their expression in placental tissues. In order to characterize the expression patterns of individual Capsicum ABC transporters at different stages (6, 16 and 25 dpa), we used publicly available RNA-seq data for C. annuum cv. CM334 [2]. The RPKM values for green fruits (16 dpa) from two varieties of C. chinense (Naga morich and Pimienta de neyde) and C. annuum cv CM334 were plotted in a hierarchical heatmap (Fig 7).

The C. annuum CM334 variety showed a similar pattern of expression in all placenta tissue stages (Fig 7A). CaABCG11 was expressed at 6 and 16 dpa with a higher expression at 6 dpa. On the other hand, CaABCB36, CaABCG83 and CaABCG87 were highly expressed at 16 dpa.
Most Capsicum ABC transporters presented different expression patterns, whereas a few resulted similar. Some exhibited stage- and species-specific expression, which suggests that these genes may play specific roles in the relevant stages and Capsicum species. Among 74 genes, 32 were expressed across all placenta tissues at different stages (Fig 7B). The ABC transporters previously described as a major marker for capsaicin and dihydrocapsaicin content were mostly expressed in C. annuum cv. CM334. CaABCC9 and CcABCC20 (CA11g01950) were found greatly expressed at 16 and 25 dpa and CaABCG28 (CA06g14430) was found in 25-dpa tissue. By contrast, CcABCG37 (CA06g14430) was greatly expressed only C. chinense cv. Naga morich at 16 dpa. Mainly, the ABCC and ABCG subfamilies were distributed across different stages; however, only ABCA, ABCB, ABCE, ABCF and ABCI members were expressed in C. annuum cv. CM334, and ABCD members were expressed in C. chinense cultivars. C. chinense varieties at 16 dpa shared the expression of eight genes (CcABCC16, CcABCC21, CcABCG45, CcABCG46, CcABCG51, CcABCG68, CcABCG74, CcABCG84). CcABCG54 was exclusively expressed in Pimienta de neyde, whereas CcABCG12, CcABCG16, CcABCG46, CcABCG51, CcABCG59, CcABCG84 and CcABCD4 were highly expressed in Naga morich. Most of the genes expressed in the C. chinense varieties belonged to the ABCC, ABCG and ABCD families.

The ABCA subfamily is not yet fully functionally characterized in plants; it has been reported to be related to pollen and seed germination and maturation [71]. The presence of one full-sized ABCA transporter was exclusive to dicots, including pepper, Arabidopsis [13], tomato [59], B. rapa [20], and B. napus [21] but so far it has not been identified in monocots, such as rice [16] and maize [18]. However, Chen et al. [58] reported one full-sized ABCA
transporter in pineapple. The ABCB subfamily is composed of a full-sized or multidrug resistance (MDR) protein and half-sized protein, with names such as transporters associated with antigen processing (TAP) and ABC transporter of mitochondria (ATM) [10]. In plants, ABCB is the second largest subfamily. For instance, in Arabidopsis, the ABCB subfamily participates in different processes such as auxin bidirectional transport, phospholipid translocation, stomatal regulation, berberine transport, Fe/S biogenesis and metal stress (Cd and Al) tolerance [72]. AtABCBl, a member of AtABCBl, has been proposed to participate in auxin transportation, and AtABCBl-overexpressing plants show long hypocotyls [73, 74].

ABCE family members are soluble ABC proteins and are also called RNase L inhibitor (RLI). They possess an N-terminal Fe-S domain, which interacts with nucleic acids [75]. Their main function has been reported to be related to control of translation and ribosome biogenesis [76]. Similarly, ABCE and ABCF family members are soluble proteins and contain an NBD-NBD domain structure. In Arabidopsis, ABCF (AtABCF3) proteins have been reported to play a role in root growth [77].

The ABCG subfamily, also called pleiotropic drug resistance or white–brown complex proteins, is the largest subfamily in plants. It has been reported that ABCGs transport various phytohormones, including abscisic acid, cytokinin, strigolactone and auxin derivatives [78]. The subcellular localization of full-sized ABCGs is the plasma membrane [79], whereas half-sized ABCGs are complex proteins and have been localized in the plasma membrane, mitochondrial membrane, chloroplast membrane and cytoplasm [18]. Full-sized ABCGs of Arabidopsis, AtABCG32 [80], and rice OsABCG31 [81] are involved principally in cuticle formation, while half-sized ABCGs play an important physiological role like cuticle formation, kanamycin resistance, abscisic acid exporter and pollen development [82–84]. In cotton, GhWBC1 a half-sized white-brown complex member has been reported to be involved in fiber cell elongation [85]. Shibata et al. [86] reported that ABCG subfamily may play a key role in export of the antimicrobial diterpene such as sesquiterpenoid phytoalexin and capsidiol for resistance to the potato late blight pathogen Phytophthora infestans in Nicotiana benthamiana.

The ABCC subfamily is also called MDR-associated proteins (MRP) because of their function in transporting glutathione- and glucuronide-conjugates in drug resistance (Verrier et al., 2008). Pang et al. [18] reported that most plant ABCCs are characterized as vacuolar localized proteins and a few of them have been reported to reside on the plasma membrane. The function of different ABCC members has been found in diverse plants; for example, Arabidopsis AtABCC5 [87], maize ZmMRP4 [88] and rice OsABCC13 [89] are implicated in phytate transport. AtABCC1 and AtABCC4 are involved in folate transport, while maize ZmMRP3 and grape VvABCC1 play an important role in anthocyanin accumulation in vacuoles [90, 91]. The high expression of the ABCC subfamily in pepper placental tissue of two species, and its principal function reported in other plants, suggest that the ABCC subfamily function in Capsicum spp. is the transport and accumulation of capsaicin in vacuoles of the placental tissue.

**Gene expression analysis**

We selected CaABCG28, CbABCG26 and CcABCG37 (orthologs of CA06g14430); as well as CaABCC9, CBABCC5 and CcABCC20 (orthologs of CA11g09150) for gene expression analysis by RT-qPCR (Fig 8). Gene expression was detected throughout all placenta stages and species analyzed. The orthologs of CA06g14430 corresponding to the ABCG family showed a similar expression pattern in the C. chinense varieties at different stages, but at 16 dpa, higher relative expression was found for the Naga morich variety and the lowest was for C. annuum cv. CM334 (Fig 8A). At 6 dpa, the expression was similar across all cultivars, but it was less in C. baccatum and C. chinense cv. Pimienta de neyde at 16 dpa. The remaining varieties showed
an expression pattern close to that of the ABCC family (Fig 8B). At 25 dpa, the highest expression was found for \textit{C. annuum} cv. CM334, followed by \textit{C. baccatum} cv. PBC81 and the lowest for the \textit{C. chinense} varieties. Different patterns in the expression between the orthologs of the Capsaicinoids markers previously mentioned suggest that the expression may be species-specific for each of the ABC subfamilies in \textit{Capsicum} species.

\section*{Conclusion}

Although the ABC transporter gene superfamily has been widely studied among extant organisms including plants, the present study is the first to report the presence of 572 putative ABC transporter proteins in the entire pepper genome sequences of three different \textit{Capsicum} species. Our results provide fundamental and exhaustive information about the pepper ABC transporters by performing a comprehensive genome-wide identification and expression patterns of these proteins family. Based on their evolutionary origin, phylogenetic analysis classified the ABC proteins into 8 main subfamilies (designated A to G, and I). Chromosomal mapping revealed that members of ABCG, ABCB and ABCC subfamilies were the most abundant genes, whereas the ABCD and ABCE subfamilies were manifested in a lesser abundance. Our results suggest that the ABC transporters, specifically the ABCC and ABCG subfamilies, interfere in capsaicin and dihydrocapsaicin content in pepper. Indirectly, these two subfamilies may be involved in the transportation of secondary metabolites such as capsaicinoids to the placenta vacuoles for their storage. Moreover, we suggest that the ABC and ABCG subfamilies play a role in the \textit{H}_2\textit{O}_2 detoxification process to reduce capsaicin degradation, specifically in the \textit{C. chinense} fruits. Our study will provide clues for further research on the evolution of the ABC transporter gene family and their influence in specific biological functions of \textit{Capsicum} fruits including plant growth, development and capsaicinoid content in pepper.

\section*{Supporting information}

S1 Table. Basic information on the ABC transporter gene family in \textit{C. annuum}. (XLSX)

S2 Table. Basic information on the ABC transporter gene family in \textit{C. baccatum}. (XLSX)

S3 Table. Basic information on the ABC transporter gene family in \textit{C. chinense}. (XLSX)
S4 Table. Conserved motifs present in ABC transporter in Capsicum species.
(XLSX)

S5 Table. Detailed information on gene ontology annotation results for Capsicum species.
(XLSX)

S6 Table. RPKM values for placenta tissues of orthologs of ABC transporter pairs of C. annuum and C. chinense.
(XLSX)

Acknowledgments
Authors acknowledge funding from USDA-NIFA grant 2016–06616 and the Department of Biotechnology, New Delhi, India for the associateship extended to Sudip Kumar Dutta, Scientist (Horticulture), ICAR RC NEH Region, Mizoram Centre, Kolasib, Mizoram-796081, India.

Author Contributions
Conceptualization: Carlos Lopez-Ortiz, Padma Nimmakayala, Umesh K. Reddy.
Data curation: Carlos Lopez-Ortiz, Purushothaman Natarajan.
Formal analysis: Carlos Lopez-Ortiz, Sudip Kumar Dutta, Purushothaman Natarajan, Yadira Peña-Garcia, Umesh K. Reddy.
Funding acquisition: Padma Nimmakayala, Umesh K. Reddy.
Investigation: Umesh K. Reddy.
Methodology: Carlos Lopez-Ortiz, Sudip Kumar Dutta, Purushothaman Natarajan, Yadira Peña-Garcia, Venkata Abburi, Thangasamy Saminathan, Umesh K. Reddy.
Project administration: Padma Nimmakayala, Umesh K. Reddy.
Resources: Padma Nimmakayala, Umesh K. Reddy.
Software: Carlos Lopez-Ortiz, Yadira Peña-Garcia, Umesh K. Reddy.
Supervision: Carlos Lopez-Ortiz, Padma Nimmakayala, Umesh K. Reddy.
Validation: Carlos Lopez-Ortiz.
Visualization: Padma Nimmakayala, Umesh K. Reddy.
Writing – original draft: Carlos Lopez-Ortiz, Sudip Kumar Dutta, Yadira Peña-Garcia, Umesh K. Reddy.
Writing – review & editing: Carlos Lopez-Ortiz, Yadira Peña-Garcia, Padma Nimmakayala, Umesh K. Reddy.

References
1. Ferrari C, Torres EAFdS. Biochemical pharmacology of functional foods and prevention of chronic diseases of aging. Biomedicine & Pharmacotherapy. 2003; 57(5–6):251–60.
2. Kim S, Park M, Yeom S-I, Kim Y-M, Lee JM, Lee H-A, et al. Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. Nature genetics. 2014; 46(3):270. https://doi.org/10.1038/ng.2877 PMID: 24441736
3. Qin C, Yu C, Shen Y, Fang X, Chen L, Min J, et al. Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum domestication and specialization. Proceedings of the National Academy of Sciences. 2014; 111(14):5135–40.
4. Perla V, Nadimi M, Reddy R, Hankins GR, Nimmakayala P, Harris RT, et al. Effect of ghost pepper on cell proliferation, apoptosis, senescence and global proteomic profile in human renal adenocarcinoma cells. PloS one. 2018; 13(10):e0206183. https://doi.org/10.1371/journal.pone.0206183 PMID: 30379886

5. Nimmakayala P, Abburi VL, Saminathan T, Alaparthi SB, Almeida A, Davenport B, et al. Genome-wide diversity and association mapping for capsaicinoids and fruit weight in *Capsicum annuum* L. Scientific reports. 2016; 6:38081. https://doi.org/10.1038/srep38081 PMID: 27901114

6. Dassa E, Bougie P. The ABC of ABCs: a phylogenetic and functional classification of ABC systems in living organisms. Research in microbiology. 2001; 152(3–4):211–29. PMID: 11421270

7. Jones P, George A. The ABC transporter structure and mechanism: perspectives on recent research. Cellular and Molecular Life Sciences CMLS. 2004; 61(6):682–99. https://doi.org/10.1007/s00018-003-3336-9 PMID: 15052411

8. Higgins CF. ABC transporters: from microorganisms to man. Annual review of cell biology. 1992; 8 (1):67–113.

9. Rea PA. Plant ATP-binding cassette transporters. Annu Rev Plant Biol. 2007; 58:347–75. https://doi.org/10.1146/annurev.arplant.57.032905.105406 PMID: 17263663

10. Verrier PJ, Bird D, Burla B, Dassa E, Forestier C, Geisler M, et al. Plant ABC proteins–a unified nomenclature and updated inventory. Trends in plant science. 2008; 13(4):151–9. https://doi.org/10.1016/j.tplants.2008.02.001 PMID: 18299247

11. Davies T, Coleman J. The *Arabidopsis thaliana* ATP-binding cassette proteins: an emerging superfamily. Plant, Cell & Environment. 2000; 23(5):431–43.

12. Wang B, Dukarevich M, Sun EI, Yen MR, Saier MH. Membrane porters of ATP-binding cassette transport systems are polyphyletic. Journal of Membrane Biology. 2009; 231(1):1. https://doi.org/10.1007/s00232-009-9200-6 PMID: 19806386

13. Ro Sánchez-Fernández, Davies TE Coleman JO, Rea PA. The *Arabidopsis thaliana* ABC protein superfamily, a complete inventory. Journal of Biological Chemistry. 2001; 276(32):30231–44. https://doi.org/10.1074/jbc.M103104200 PMID: 11346655

14. Martinoia E, Klein M, Geisler M, Bovet L, Forestier C, Kolukisaoglu Ü, et al. Multifunctionality of plant ABC transporters–more than just detoxifiers. Planta. 2002; 214(3):345–55. PMID: 11855639

15. Moon S, Jung K-H. Genome-wide expression analysis of rice ABC transporter family across spatio-temporal samples and in response to abiotic stresses. Journal of plant physiology. 2014; 171(14):1276–88. https://doi.org/10.1016/j.jplph.2014.05.006 PMID: 25014263

16. Saha J, Sengupta A, Gupta K, Gupta B. Molecular phylogenetic study and expression analysis of ATP-binding cassette transporter gene family in *Oryza sativa* in response to salt stress. Computational biology and chemistry. 2015; 54:18–32. https://doi.org/10.1016/j.compbiolchem.2014.11.005 PMID: 25531538

17. Lane TS, Rempe CS, Davitt J, Staton ME, Peng Y, Soltis DE, et al. Diversity of ABC transporter genes across the plant kingdom and their potential utility in biotechnology. BMC biotechnology. 2016; 16 (1):47. https://doi.org/10.1186/s12896-016-0277-6 PMID: 27245738

18. Pang K, Li Y, Liu M, Meng Z, Yu Y. Inventory and general analysis of the ATP-binding cassette (ABC) gene superfamily in maize (*Zea mays* L.). Gene. 2013; 526(2):411–28. https://doi.org/10.1016/j.gene.2013.05.051 PMID: 23743399

19. Chen L, Li Y-x, Li C, Shi Y, Song Y, Zhang D, et al. Genome-wide analysis of the pentatricopeptide repeat gene family in different maize genomes and its important role in kernel development. BMC plant biology. 2018; 18(1):366. https://doi.org/10.1186/s12870-018-1572-2 PMID: 30567489

20. Yan C, Duan W, Lyu S, Li Y, Hou X. Genome-wide identification, evolution, and expression analysis of the ATP-binding cassette transporter gene family in *Brassica rapa*. Frontiers in plant science. 2017; 8:349. https://doi.org/10.3389/fpls.2017.00349 PMID: 28367152

21. Zhang XD, Zhao KK, Yang ZM. Identification of genomic ATP binding cassette (ABC) transporter genes and Cd-responsive ABCs in *Brassica napus*. Gene. 2018; 664:139–51. https://doi.org/10.1016/j.gene.2018.04.060 PMID: 29709635

22. Çakır B, Kilçkaya O. Whole-genome survey of the putative ATP-binding cassette transporter family genes in *Vitis vinifera*. PloS one. 2013; 8(11):e78860. https://doi.org/10.1371/journal.pone.0078860 PMID: 24244377

23. Theodoulou FL, Holdsworth M, Baker A. Peroxisomal ABC transporters. Febs Letters. 2006; 580 (4):1139–55. https://doi.org/10.1016/j.febslet.2005.12.095 PMID: 16413537

24. Ito H, Gray WM. A gain-of-function mutation in the *Arabidopsis* pleiotropic drug resistance transporter PDR9 confers resistance to auxinic herbicides. Plant physiology. 2006; 142(1):63–74. https://doi.org/10.1104/pp.106.084533 PMID: 16877699
Kang J, Park J, Choi H, Burla B, Kretzschmar T, Lee Y, et al. Plant ABC transporters. The Arabidopsis book/American Society of Plant Biologists. 2011; 9.

Jing H, Li C, Ma F, Ma J-H, Khan A, Wang X, et al. Genome-wide identification, expression diversification of dehydrin gene family and characterization of CaDHNN3 in pepper (Capsicum annuum). PLoS one. 2016; 11(8):e0161073. https://doi.org/10.1371/journal.pone.0161073 PMID: 27551973

ZHANG H, Ning C, Chunjuan D, SHANG Q. Genome-wide Identification and Expression of ARF Gene Family during Adventitious Root Development in Hot Pepper (Capsicum annuum). Horticultural plant journal. 2017; 3(4):151–64.

Khan A, Li R-J, Sun J-T, Ma F, Zhang H-X, Jin J-H, et al. Genome-wide analysis of dirigent gene family in pepper (Capsicum annuum L.) and characterization of CaDIR7 in biotic and abiotic stresses. Scientific reports. 2018; 8(1):5000. https://doi.org/10.1038/s41598-018-23761-0 PMID: 29615685

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. Nucleic acids research. 2011; 40(D1):D1178–D86.

Eddy SR. Profile hidden Markov models. Bioinformatics (Oxford, England). 1998; 14(9):755–63.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic acids research. 2013; 42(D1):D222–D30.

Letunic I, Bork P. 20 years of the SMART protein domain annotation resource. Nucleic acids research. 2017; 46(D1):D493–D6.

Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. Nucleic acids research. 2018; 46(W1):W200–W4. https://doi.org/10.1093/nanogr/gyy446 PMID: 29905871

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular biology and evolution. 2018; 35(6):1547–9. https://doi.org/10.1093/molbev/msy096 PMID: 29722887

Thompson JD, Gibson TJ, Pilewski F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic acids research. 1997; 25(24):4876–82. PMID: 9396791

Letunic I, Bork P. Interactive Tree Of Life (iTOl): an online tool for phylogenetic tree display and annotation. Bioinformatics. 2006; 23(1):127–8. https://doi.org/10.1093/bioinformatics/btl529 PMID: 17050570

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic acids research. 2003; 31(13):3784–8. PMID: 12824418

Higo K, Ugawa Y, Iwamoto M, Korenaga T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic acids research. 1999; 27(1):297–300. PMID: 9847208

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clement L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic acids research. 2009; 37(suppl_2):W202–W8.

De Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic acids research. 2006; 34(suppl_2):W362–W5.

Gu X, Wang Y, Gu J. Age distribution of human gene families shows significant roles of both large-and small-scale duplications in vertebrate evolution. Nature genetics. 2002; 31(2):205. https://doi.org/10.1038/ng902 PMID: 12032571

Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genomics, proteomics & bioinformatics. 2010; 8(1):77–80.

Moniz de Sa M, Drouin G. Phylogeny and substitution rates of angiosperm actin genes. Molecular biology and evolution. 1996; 13(9):1198–212. https://doi.org/10.1093/oxfordjournals.molbev.a025685 PMID: 8896372

Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet journal. 2011; 17(1):10–2.

Joshi N, Fass J. Sickle-A windowed adaptive trimming tool for FASTQ files using quality. Online publication https://github.com/najoshi/sickle[Google Scholar]. 2011.

Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics. 2010; 26(5):589–95. https://doi.org/10.1093/bioinformatics/btp698 PMID: 20080505

Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015; 31(2):166–9. https://doi.org/10.1093/bioinformatics/btu638 PMID: 25260700

Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic acids research. 2015; 43(W1):W566–W70. https://doi.org/10.1093/nar/gkv468 PMID: 25969447
49. Fernandez-Pozo N, Menda N, Edwards JD, Saha S, Teclie Y, Strickler SR, et al. The Sol Genomics Network (SGN)—from genotype to phenotype to breeding. Nucleic acids research. 2014; 43(D1): D1036–D41.

50. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. methods. 2001; 25(4):402–8. https://doi.org/10.1006/meth.2001.1262 PMID: 11846609

51. Wan H, Yuan W, Ruan M, Ye Q, Wang R, Li Z, et al. Identification of reference genes for reverse transcription quantitative real-time PCR normalization in pepper (Capsicum annuum L.). Biochemical and biophysical research communications. 2011; 416(1–2):24–30. https://doi.org/10.1016/j.bbrc.2011.10.105 PMID: 22086175

52. Hoffman PG, Lego MC, Galetto WG. Separation and quantitation of red pepper major heat principles by reverse-phase high-pressure liquid chromatography. Journal of Agricultural and Food Chemistry. 1983; 31(6):1326–30.

53. Govindarajan V, Sathyanarayana M. Capsicum—production, technology, chemistry, and quality. Part V. Impact on physiology, pharmacology, nutrition, and metabolism; structure, pungency, pain, and desensitization sequences. Critical Reviews in Food Science & Nutrition. 1991; 29(6):435–74.

54. Sanatombi K, Sharma G. In vitro propagation of Capsicum chinense Jacq. Biologia Plantarum. 2008; 52(3):517–20.

55. García-Claver A, Fellman SM, Gil-Ortega R, Jahn M, Arnedo-Anderes MS. Identification, validation and survey of a single nucleotide polymorphism (SNP) associated with pungency in Capsicum spp. Theoretical and Applied Genetics. 2007; 115(7):907–16. https://doi.org/10.1007/s00122-007-0617-y PMID: 17882396

56. Ofori PA, Mizuno A, Suzuki M, Martinoia E, Reuscher S, Aoki K, et al. Genome-wide analysis of ATP binding cassette (ABC) transporters in tomato. PloS one. 2018; 13(7):e0200854. https://doi.org/10.1371/journal.pone.0200854 PMID: 30048467

57. Linton KJ, Higgins CF. Structure and function of ABC transporters: the ATP switch provides flexible control. Pflligers Archiv-European Journal of Physiology. 2007; 453(5):555–67. https://doi.org/10.1007/s00424-006-0126-x PMID: 16937116

58. Li Y, Chen X, Chen Z, Cai R, Zhang H, Xiang Y. Identification and expression analysis of BURP domain-containing genes in Medicago truncatula. Frontiers in plant science. 2016; 7:485. https://doi.org/10.3389/fpls.2016.00426 PMID: 27148311

59. Paul A, Rao S, Mathur S. The α-crystallin domain containing genes: identification, phylogeny and expression profiling in abiotic stress, phytohormone response and development in tomato (Solanum lycopersicum). Frontiers in plant science. 2016; 7:426. https://doi.org/10.3389/fpls.2016.00426 PMID: 27066058

60. Lee JM, Kim S, Lee JY, Yoo EY, Cho MC, Cho MR, et al. A differentially expressed proteomic analysis in placental tissues in relation to pungency during the pepper fruit development. Proteomics. 2006; 6(19):5248–59. https://doi.org/10.1002/pmic.200600326 PMID: 16947123

61. Naves ER, de Ávila Silva L, Sulpice R, Araújo WL, Nunes-Nesi A, Peres LE, et al. Capsaicinoids: Pungency beyond Capsicum. Trends in plant science. 2018.

62. Bernal MA, Calderon AA, Pedreno MA, Munoz R, Ros Barceló A, Merino de Caceres F. Capsaicin oxidation by peroxidase from Capsicum annum (variety Annuum) fruits. Journal of Agricultural and Food Chemistry. 1993; 41(7):1041–4.

63. Díaz J, Pomar F, Bernal A, Merino F. Peroxidases and the metabolism of capsaicin in Capsicum annum. Phytochemistry Reviews. 2004; 3(1–2):141–57.
68. Johnson CD, Decoteau DR. Nitrogen and potassium fertility affects jalapeño pepper plant growth, pod yield, and pungency. HortScience. 1996; 31(7):1119–23.

69. Medina-Lara F, Echevarría-Machado I, Pacheco-Arjona R, Ruiz-Lau N, Guzmán-Antonio A, Martínez-Estevez M. Influence of nitrogen and potassium fertilization on fruiting and capsacain content in habanero pepper (Capsicum chinense Jacq.). HortScience. 2008; 43(5):1549–54.

70. Woldemariam SH, Lai S, Zelelew DZ, Solomon MT. Effect of Potassium Levels on Productivity and Fruit Quality of Tomato (Lycopersicon esculentum L.). 2018. 2018; 6(1):14. Epub 2017-12-10.

71. Garcia O, Bougie P, Forestier C, Dassa E. Inventory and comparative analysis of rice and Arabidopsis ATP-binding cassette (ABC) systems. Journal of molecular biology. 2004; 343(1):249–65. https://doi.org/10.1016/j.jmb.2004.07.093 PMID: 15381434

72. Hwang J-U, Song W-Y, Hong D, Ko D, Yamaoka Y, Jang S, et al. Plant ABC transporters enable many unique aspects of a terrestrial plant’s lifestyle. Molecular Plant. 2016; 9(3):338–55. https://doi.org/10.1016/j.molp.2016.02.003 PMID: 26902186

73. Sidler M, Hassa P, Hasan S, Ringli C, Dudler R. Involvement of an ABC transporter in a developmental pathway regulating hypocotyl cell elongation in the light. The Plant Cell. 1998; 10(10):1623–36. PMID: 9761790

74. Noh B, Murphy AS, Spalding EP. Multidrug resistance–like genes of Arabidopsis required for auxin transport and auxin-mediated development. The Plant Cell. 2001; 13(11):2441–54. https://doi.org/10.1105/tpc.010350 PMID: 11701880

75. Andolfog R, Ruocco M, Di Donato A, Frusciante L, Lorito M, Scala F, et al. Genetic variability and evolutionary diversification of membrane ABC transporters in plants. BMC plant biology. 2015; 15(1):51.

76. Dong J, Lai R, Nielsen K, Fekete CA, Qiu H, Hinnebusch AG. The essential ATP-binding cassette protein RLI1 functions in translation by promoting preinitiation complex assembly. Journal of Biological Chemistry. 2004; 279(40):42157–68. https://doi.org/10.1074/jbc.M404502200 PMID: 15277527

77. Kato T, Tabata S, Sato S. Analyses of expression and phenotypes of knockout lines for Arabidopsis ABCF subfamily members. Plant biotechnology. 2009; 26(4):409–14.

78. Borghi L, Kang J, Ko D, Lee Y, Martinoia E. The role of ABCG-type ABC transporters in phytohormone transport. Biochemical Society Transactions. 2015; 43(5):924–30. https://doi.org/10.1042/BST20150106 PMID: 26517905

79. Banasiak J, Jasinski M. Defence, symbiosis and ABCG transporters. Plant ABC Transporters: Springer; 2014. p. 163–84.

80. Bessire M, Borel S, Fabre G, Carraça L, Efremova N, Yephremov A, et al. A member of the PLEIOTROPIC DRUG RESISTANCE family of ATP binding cassette transporters is required for the formation of a functional cuticle in Arabidopsis. The Plant Cell. 2011; 23(5):1958–70. https://doi.org/10.1105/tpc.111.083121 PMID: 21628525

81. Chen G, Komatsuda T, Ma JF, Nawrath C, Pourkheirandish M, Tagiri A, et al. An ATP-binding cassette subfamily G full transporter is essential for the retention of leaf water in both wild barley and rice. Proceedings of the National Academy of Sciences. 2011; 108(30):12354–9.

82. Mentewab A, Stewart CN Jr. Overexpression of an Arabidopsis thaliana ABC transporter confers kanamycin resistance to transgenic plants. Nature biotechnology. 2005; 23(9):1177. https://doi.org/10.1038/nbt1134 PMID: 16116418

83. Kuromori T, Ito T, Sugimoto E, Shinozaki K. Arabidopsis mutant of ATABCG26, an ABC transporter gene, is defective in pollen maturation. Journal of plant physiology. 2011; 168(16):2001–5. https://doi.org/10.1016/j.jplph.2011.05.014 PMID: 21696448

84. Xu XM, Møller SG. AtNAP7 is a plastidic SufC-like ATP-binding cassette/ATPase essential for Arabidopsis embryogenesis. Proceedings of the National Academy of Sciences. 2004; 101(24):9143–8.

85. Zhu Y-Q, Xu K-X, Luo B, Wang J-W, Chen X-Y. An ATP-binding cassette transporter GhWBC1 from elongating cotton fibers. Plant physiology. 2003; 133(2):580–8. https://doi.org/10.1104/pp.103.027052 PMID: 12972649

86. Shibata Y, Ojika M, Sugiyama A, Yazaki K, Jones DA, Kawakita K, et al. The full-size ABCG transporters Nb-ABC G1 and Nb-ABC G2 function in pre-and postinvasion defense against Phytophthora infestans in Nicotiana benthamiana. The Plant Cell. 2016; 28(5):1163–81. https://doi.org/10.1105/tpc.15.00721 PMID: 27102667

87. Nagy R, Grob H, Weder B, Green P, Klein M, Frelet-Barrand A, et al. The Arabidopsis ATP-binding cassette protein AtMRP5/AtABC5 is a high affinity inositol hexakisphosphate transporter involved in guard cell signaling and phytate storage. Journal of Biological Chemistry. 2009; 284(48):33614–22. https://doi.org/10.1074/jbc.M109.030247 PMID: 19797057
88. Badone FC, Cassani E, Landoni M, Doria E, Panzeri D, Lago C, et al. The low phytic acid1-241 (lpa1-241) maize mutation alters the accumulation of anthocyanin pigment in the kernel. Planta. 2010; 231 (5):1189–99. https://doi.org/10.1007/s00425-010-1123-z PMID: 20191364

89. Tagashira Y, Shimizu T, Miyamoto M, Nishida S, Yoshida K. Overexpression of a gene involved in phytic acid biosynthesis substantially increases phytic acid and total phosphorus in rice seeds. Plants. 2015; 4(2):196–208. https://doi.org/10.3390/plants4020196 PMID: 27135323

90. Goodman CD, Casati P, Walbot V. A multidrug resistance–associated protein involved in anthocyanin transport in Zea mays. The Plant Cell. 2004; 16(7):1812–26. https://doi.org/10.1105/tpc.022574 PMID: 15208386

91. Francisco RM, Regalado A, Ageorges A, Burla BJ, Bassin B, Eisenach C, et al. ABCC1, an ATP binding cassette protein from grape berry, transports anthocyanidin 3-O-glucosides. The Plant Cell. 2013; 25 (5):1840–54. https://doi.org/10.1105/tpc.112.102152 PMID: 23723325