RESEARCH

Origin, evolution and diversification of plant mechanosensitive channel of small conductance-like (MSL) proteins

Zaibao Zhang[1*†], Fan Ye[2†], Tao Xiong[3], Jiahui Chen[2], Jiajia Cao[2], Yurui Chen[2] and Sushuang Liu[1]

Abstract

Mechanosensitive (MS) ion channels provide efficient molecular mechanism for transducing mechanical forces into intracellular ion fluxes in all kingdoms of life. The mechanosensitive channel of small conductance (MscS) was one of the best-studied MS channels and its homologs (MSL, MscS-like) were widely distributed in cell-walled organisms. However, the origin, evolution and expansion of MSL proteins in plants are still not clear. Here, we identified more than 2100 MSL proteins from 176 plants and conducted a broad-scale phylogenetic analysis. The phylogenetic tree showed that plant MSL proteins were divided into three groups (I, II and III) prior to the emergence of chlorophytae algae, consistent with their specific subcellular localization. MSL proteins were distributed unevenly into each of plant species, and four parallel expansion was identified in angiosperms. In Brassicaceae, most MSL duplicates were derived by whole-genome duplication (WGD)/segmental duplications. Finally, a hypothetical evolutionary model of MSL proteins in plants was proposed based on phylogeny. Our studies illustrate the evolutionary history of the MSL proteins and provide a guide for future functional diversity analyses of these proteins in plants.

Keywords Mechanosensitive ion channel, MscS-like (MSL), Molecular evolution, Origin, Expansion

Introduction

All living organisms are subjected to various external and internal mechanical stresses, including gravity, touch, sound and osmotic shock. How mechanical forces are sensed by cells is a long-standing question in biology. One of the most universal mechanisms for cells to respond to mechanical stimuli is the use of mechanosensitive (MS) ion channels [1]. MS channels are transmembrane proteins that exist in all kingdoms of life. The primary function of MS channels is to provide a conductive pore in response to mechanical stimulation, allowing ions to flow across the membrane down their electrochemical gradient [2].

Many types of MS ion channels have been identified in different organisms, including mechanosensitive ion channel of small conductance (MscS) [3, 4], mechanosensitive ion channel of large conductance (MscL) [5], two pore potassium (TPK) [6], Mid1-complementing activity (MCA) [7], and piezo [8]. Different MS channels displayed highly divergent in conductance, ion selectivity, and/or sensitivity to the direction of activation pressure. MscS is a nonselective stretch-activated channel which is gated by membrane tension [9, 10]. *Escherichia coli* has six MscS paralogs: archetypal MscS (yggB), potassium-dependent MscK (kefA), MscM (YjeP), YbdG, YbiO and YnaI [11, 12]. These bacteria MscS channels have different

†Zaibao Zhang and Fan Ye contributed equally to this work.

*Correspondence:
Zaibao Zhang
zaibaozhang79@163.com
[1] School of Life and Health Science, Huzhou College, Huzhou, Zhejiang, China
[2] College of International Education, Xinyang Normal University, Xinyang, Henan, China
[3] College of Life Science, Xinyang Normal University, Xinyang, Henan, China

Zhang *et al. BMC Plant Biology*   (2023) 23:462

Page 2 of 14

activation thresholds and channel conductance, protecting cells from osmotic stress by providing a conduit for the release of osmolytes from the bacterium [13–15].

MscS family members are highly divergent in their topology and domain structure. The crystal structure of *E. coli* archetypal MscS (PDB: 2OAU – EcMscS, PDB) was resolved [16], and it was characterized by three N-terminal transmembrane (TM) helices followed by a large hydrophilic cytoplasmic domain. The key feature of the EcMscS structure is the pore lining TM helix, TM3, which forms a hydrophobic channel pore and shows the highest homology in MscS-like channels [4, 17]. TM3 comprises two regions, TM3a and TM3b, which are separated by a distinctive kink at residue G113. A comparison of the open-state versus closed-state structures of EcMscS showed that gating involves swinging a tension-sensitive paddle made up of the TM1/TM2 helices and twisting TM3a at G113 [18, 19].

MscS homologs are widely dispersed in bacterial, archaeal, fungal and plant genomes, but not in animal genomes. In Arabidopsis, ten MscS homologs were identified, named as MscS-like proteins (MSLs) [20]. The Arabidopsis MSLs were divided into three phylogenetic groups, consist with their different subcellular localization and topology [21]. Group I (AtMSL1) and group II MSLs (AtMSL2 and AtMSL3) were localized in the inner membrane of mitochondria and chloroplast, respectively [20, 22, 23]. Group III MSLs (AtMSL4-10) were localized in the plasma or endoplasmic reticulum (ER) membrane [20, 21, 24]. Both group I and group II MSLs contained five TM helices, while group III MSLs contained six TM helices. The distinct cellular localizations implicate diverse physiological functions of plant MSLs. Loss of *AtMSL1* increased mitochondrial oxidation under abiotic stresses, indicating that AtMSL1 is crucial for regulating mitochondrial redox status under abiotic stress [25]. AtMSL2 and AtMSL3 colocalize with the plastid division protein AtMinE and function redundantly in maintaining plastid shape, size and division [22, 26, 27]. *AtMSL8* is specially expressed in pollen, and plays essential roles in pollen hydration and pollen tube growth during fertilization [28, 29]. Both loss of function and overexpression of *AtMSL8* lead to reduced pollen germination and low fertility. *AtMSL10* is expressed in root and form a heteromeric channel with AtMSL9 [23, 30]. Both loss of function and overexpression of *AtMSL10* lead to growth retardation and ectopic cell death [31]. AtMSL10 is also involved in the wound-triggered early signal transduction and plays a positive regulatory role in biosynthesis of jasmonic acid [32]. Moreover, the MSLs were identified and analyzed in *Oryza sativa, Aegilops tauschii, Hordeum vulgare, Sorghum bicolor, Triticum aestivum, Triticum urartu, Zea mays, Phaseolus vulgaris*, and *Cicer arietinum*, respectively [33–38]. The majority of *MSL* genes in these plants were expressed in various tissue/organs. In rice, most *MSL* genes were significantly expressed in reproductive stages [33]. In maize, four *MSL* genes were expressed in all the tissues development stages, and one maize *MSL* gene specifically expressed in reproductive tissues [37]. The differential expression patterns of *MSLs* indicating their functions in different tissues and organs.

The MscS homologs were discovered in Bacteria, suggesting that MscS existed in the early stages of evolution. However, its identification and functional analyses in plants are still limited. In order to explore the origin, characterization and diversification of plant MSLs, we sought to build a comprehensive phylogeny of plant MSLs. 2123 MSL proteins were identified from 176 plants. Based on the phylogenetic tree, we explored the origin and divergence of the plant MSL proteins. Three MSL groups (I, II and III) were identified in plants, and the divergence of these three MSL groups can be traced back to chlorophytae algae. In addition, a wide phylogenetic architecture of angiosperm MSLs were constructed, and the MSLs in angiosperms were further classified to 4 clades: MSL1, MSL2/3, MSL4-8 and MLS9/10. Finally, we discussed the possible evolutionary relationships of the MSL proteins in plants.

Results
MSL proteins were identified in all lineages of plant
By employing the Hidden Markov Models (HMM) algorithm and BLASTP search, we constructed a comprehensive phylogeny of MSL proteins in genome-sequenced plants by identifying MSL proteins in genomes of 2 chlorophytic algaes, 6 charophytes, 5 bryophytes, 2 ferns, 5 gymnosperms and 156 angiosperms (Supplementary Table 1). The retrieved proteins were checked with SMART and PFAM, and the candidates containing the MS_channel domain were regarded as 'true' MSLs. In total, 2123 MSL proteins were retrieved from 176 plant species (Table 1). The copy number of MSL proteins varies among different plant lineages, ranging on average from 5.5 copies in chlorophyta, 4.7 in bryophytes, 9.0 in ferns, 6.0 in gymnosperms, to 12.7 angiosperms (Table 1, Supplementary Table 2). These data suggest that the MSL proteins were expanded in angiosperms. Among the angiosperms examined, more MSL proteins were identified in dicots than that in monocots, with an average of 9.2 in Poaceae, 12.5 in Brassicaceae, 12.0 in Leguminosae and 11.1 in Rosaceae (Table 1, Supplementary Table 2). These results indicate that MSL proteins were expanded in dicots.

Diversification of MSL proteins in green plants
MSL proteins are distributed in bacterial, archaeal and plant genomes, however, they are not found in animal genomes. The full length of MSLs exhibit little similarity, however, the hydrophobic pore-lining TM3 helix and upper cytoplasmic domain display high similarity. Using the pore-lining

Zhang *et al. BMC Plant Biology*   (2023) 23:462

Page 3 of 14

Table 1 The number of MSL proteins in green plants

Taxonomy	Number of species	Number of MSL	Average number of MSL per species
Chlorophytes	2	11	5.50
Charophytes	6	28	4.67
Bryophytes	5	49	9.80
Ferns	2	18	9.00
Gymnosperms	5	30	6.00
Angiosperms	156	1971	12.71
Poaceae	22	204	9.27
Brassicaceae	10	125	12.50
Leguminosae	13	156	12.00
Rosaceae	9	100	11.11

helix and the conserved cytoplasmic regions of MSLs, a predicted evolutionary tree was constructed among representative members of MSL homologs from bacteria, fungi, protozoa and plants (Figure S1). The plant MSL proteins are separated from bacteria, protozoa and fungi MSLs, and fall into three distinct phylogenetic groups. These results indicated that the divergence of plant MSLs occurred after the emergence of plants. To further explore the evolutionary origin of MSL proteins in green plants, we reconstructed phylogenetic tree with MSL homologs from 16 representative viridiplantae species including green algae *Chlamydomonas reinhardtii*, moss *Physcomitrella patens*, fern *Selaginella moellendorffii*, gymnosperm *Pinus lambertiana*, basal angiosperm *Amborella trichopoda*, monocots (*Oryza sativa*, *Zea mays*), and eudicots (*Arabidopsis thaliana*, *Arabidopsis lyrata*, *Theobroma cacao*, *Glycine max*, *Camellia sinensis*, *Coffea canephora*) (Fig. 1). The topology of the phylogenetic tree clearly divided plant MSLs into three clades (Clade I, II, III) (Fig. 1a), consistent with the three distinct subcellular localizations predictions (Fig. 1b). Each clade contains genes from several major lineages of green plants, including algae, mosses, and gymnosperms, indicating that plant MSLs originated in the ancestors of green plants. The angiosperm clade I (MSL1) is monophyletic group. In clade II (MSL2/3), MSL2 and MSL3 diverged in eudicots, demonstrating the diversification of MSL2 and MSL3 occurred in the ancestor of eudicots. More MSL homologs are identified in clade III (MSL4-10). A distinction between MSL4-8 and MSL9/10 are observed in seed plants (gymnosperms and angiosperms), suggesting that the diversification into MSL4-8 and MSL9/10 occurred in the ancestor of seed plants.

## Phylogenetic classification of the MSL proteins in angiosperms

To further explore the phylogenetic relationship of MSL proteins in angiosperms, we reconstructed a wide phylogenetic tree with 1971 MSL proteins identified from 155 angiosperm species (Fig. 2). The phylogenetic tree shows that angiosperm MSLs were divided into four major groups (MSL1, MSL2/3, MSL4-8, MSL9/10) (Fig. 2). Many species-specific amplifications of MSL1 are identified in angiosperm, with 6, 5, 7, 13, and 7 MSL1 homologs are identified in rice, *Brachypodium distachyon*, *Phoenix dactylifera*, *Dendrobium catenatum*, *Nicotiana tabacum*, respectively (Fig. 2). MSL2/3 was divided into MSL2 and MLS3 in eudicots, and most plants have more MSL homologs corresponding to Arabidopsis MSL2 and MLS3 (Figs. 2 and 3). MSL4-8 and MSL9/10 are lineage-specific paralogs within Brassicaceae (Figs. 2, 4 and 5). In addition, compared with monocots, more MSL homologs were identified in eudicots, indicating that MSLs were expanded in eudicot. The topological structure of the gene tree is similar to that of species tree, indicating that these four MSL clades originated independently.

In the MSL2/3 clade, MSL2 and MSL3 shared common dicot lineages and no monocot and magnoliidae species were included, indicating that the divergence between MSL2 and MSL3 occurred before the emergence of dicots and after the monocots/magnoliidae/dicots divergence (Fig. 3). In the MSL4/5/6/7/8 (MSL4-8) clade, two sub-branches were identified in dicots: Dicots-Group I and Dicots-Group II (Fig. 4). MSL4/5/6/7/8 of all Brassicaceae plants were clustered in Dicots-Group I, and Brassicales MSLs were not appear in Dicots-Group II. In addition, MSL4-8 in Brassicaceae experienced at least three expansion events, resulting in 5 copies of MSL in each species (Fig. 4). In the MSL9/10 clade, three sub-branches were identified in dicots: Dicots-Group I, Dicots-Group II and Dicots-Group III (Fig. 5). MSL9/10 of all Brassicaceae plants were clustered in Dicots-Group I, and Brassicaceae MSLs were lost in Dicots-Group II and III. The differentiation between MSL9 and MSL10 is due to Brassicaceae-specific duplication (Fig. 6).

Gene structure, conserved motif and phylogenetic analysis of MSLs in Brassicaceae

MSLs were expanded in Brassicaceae. To explore the expansion of MSLs in Brassicaceae, we constructed a phylogenetic tree using Database III, which includes 10 Brassicaceae plants and 1 Brassicales plant (*Carica papaya*) (Fig. 6). The phylogenetic tree showed that MSL4-8 and MSL9/10 are expanded in Brassicaceae, consistent with the former phylogenetic analysis (Fig. 6). The number of exon/intron in Brassicaceae MSLs was analyzed (Fig. 6). In Group I, most MSLs (4/7) have 5 introns. In Group II, most MSLs (12/15) have 12 introns. In Group III, more than two thirds of MSLs (51/64) contained 4 introns. Therefore, MSL members belong to different groups displayed different exon/intron structures, while MSL members belong to the same group showed similar exon/intron distribution (Supplementary Table 3).

Zhang *et al. BMC Plant Biology*   (2023) 23:462

Page 4 of 14
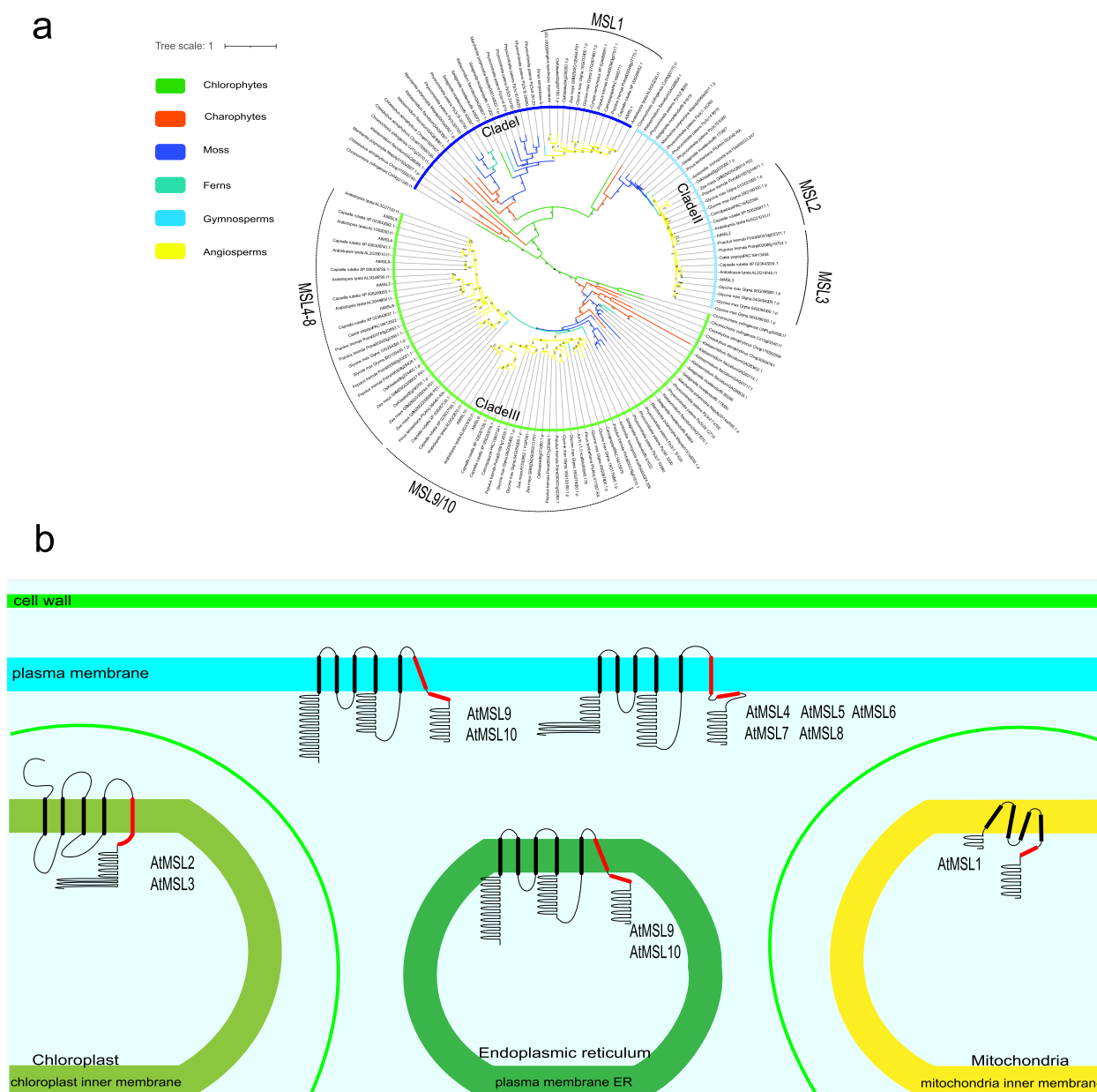
Fig. 1 Phylogenetic relationships and subcellular localizations of MSL proteins. **a** An phylogenetic tree of MSL proteins using Bayesian method in green plants. 187 MSL proteins from 3 chlorophytes, 13 charophytes, 5 bryophytes, 1 lycophyte, 2 ferns, 5 gymnosperms and 4 angiosperms were included in phylogeny. **b** Predicted subcellular localization and topology of MSL proteins from *Arabidopsis thaliana* (modify from Hamilton et al., 2015 and Li et al., 2020) [21, 39]. Topologies were drawn according to predictions on WoLF PSORT (https://wolfpsort.hgc.jp). The regions of highest homology to *E.Coli* MscS TM3 were shown in red

The conserved motif was predicted by the MEME tool (Fig. 6). MSL members belonging to Group I only have 1 (motif 3) or 2 motifs (motif 3, and motif 6). Most MSL members in Group II and Group III have 6 motifs. However, the motif locations are different between Group II and Group III MSLs. Both gene structure and protein motif results showed that MSL genes in the same group had similar gene structure and motif composition, indicating that there have

similar functions. In addition, MSL genes between different groups display significant difference in both gene structure and motif composition, indicating the different functions.

Expansion pattern of *MSLs* genes

Whole genome/segment and tandem duplications contribute significantly to the expansion of gene families, and gene duplication promotes genome evolution [40].
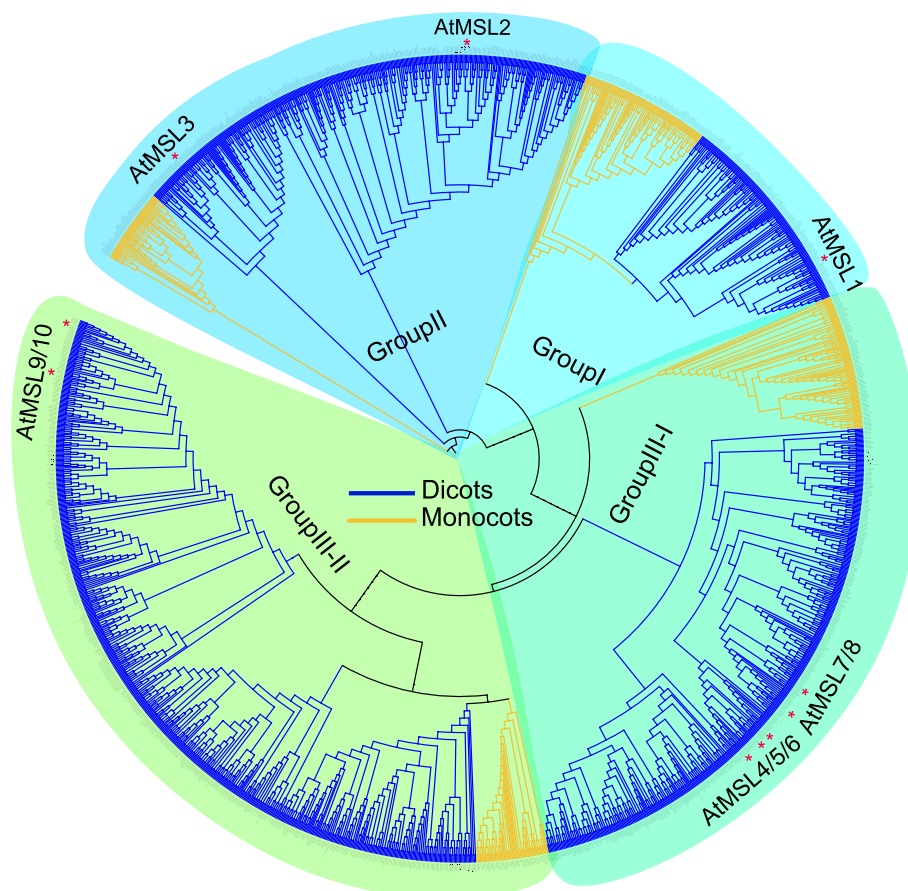
Fig. 2 Phylogenetic classification of MSL proteins in angiosperms. The topology shows that MSLs in angiosperms can be classified into 4 sub-groups: MSL1, MSL2/3, MSL4-8, and MSL9/10. Dicots are marked in blue and monocots are marked in yellow

To explore the expansion of *MSL* genes, we conducted a comprehensive synteny analysis (Fig. 7). The results showed that the Arabidopsis *MSL* genes had 3 segmental duplication events (*AtMSL4/AtMSL5, AtMSL4/AtMSL7, AtMSL9/AtMSL10*), and one tandem duplication event (*AtMSL7/AtMSL8*). In addition, all these expanded *MSL* genes belongs to group III of MSL (Fig. 7). Similar expansion patterns were also identified in other Brassicaceae plants, including *Arabidopsis lyrate, Brassica niqra*, and *Brassica oleracea* (Fig. 7).

Furthermore, we performed a multicollinearity analysis of *MSL* genes among 6 Brassicales species (*Carica papaya, Arabidopsis thaliana, Arabidopsis lyrate, Brassica niqra, Brassica oleracea,* and *Capsella grandiflora*) to reveal the robust orthologs of these *MSLs* (Fig. 7). The results showed that among Brassicaceae species, all *MSL* genes were collinearity. In addition, a relatively low collinearity was found between *Carica papaya* and other Brassicaceae species. These results indicate that the *MSL* genes were conserved and have the same ancestors.

Discussion

In this study, we have performed a comprehensive evolutionary analysis of the *MSL* gene family in green plants. The phylogenetic insights provide valuable information for future molecular and biological investigations of various MSL proteins.

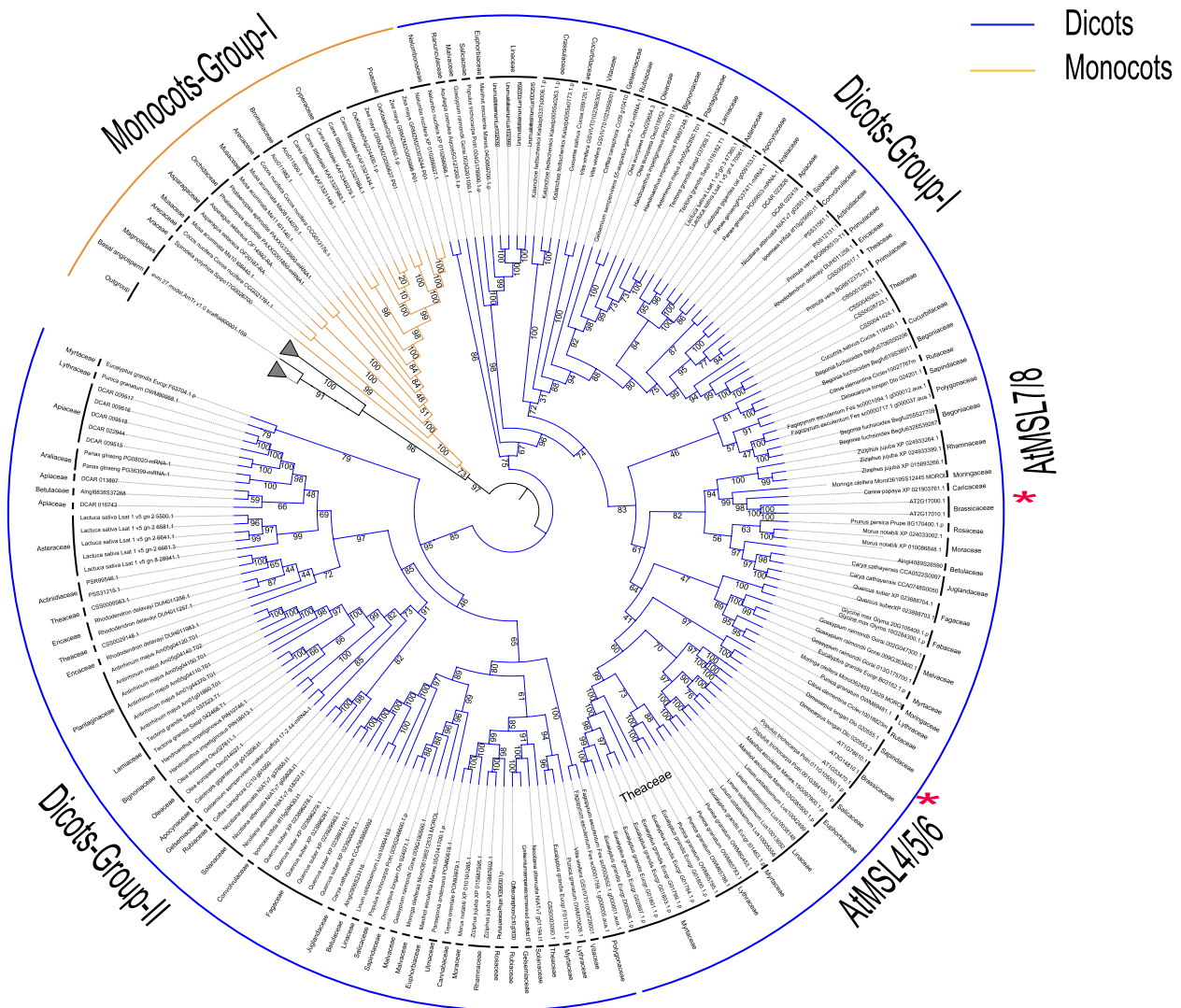### Phylogenetic relationship of plant MSLs

Mechanosensitive (MS) ion channels provide molecular mechanism for the cellular response to mechanical stimuli, and are widely identified in bacteria, plants, animals and humans. Mechanosensitive channel of small conductance (MscS) is one of the best-studied MS channels, and MscS homologs are widely dispersed among the bacterial, fungi and plant lineages. However, MscS homologs have not been identified in animals, indicating that they can serve as therapeutic targets for pathogenic bacteria, fungi and protozoa. Our search showed that MSL proteins are widely exist in green lineage of plants and the copy number of MSL is varies among different

Zhang *et al. BMC Plant Biology*   (2023) 23:462

Page 6 of 14
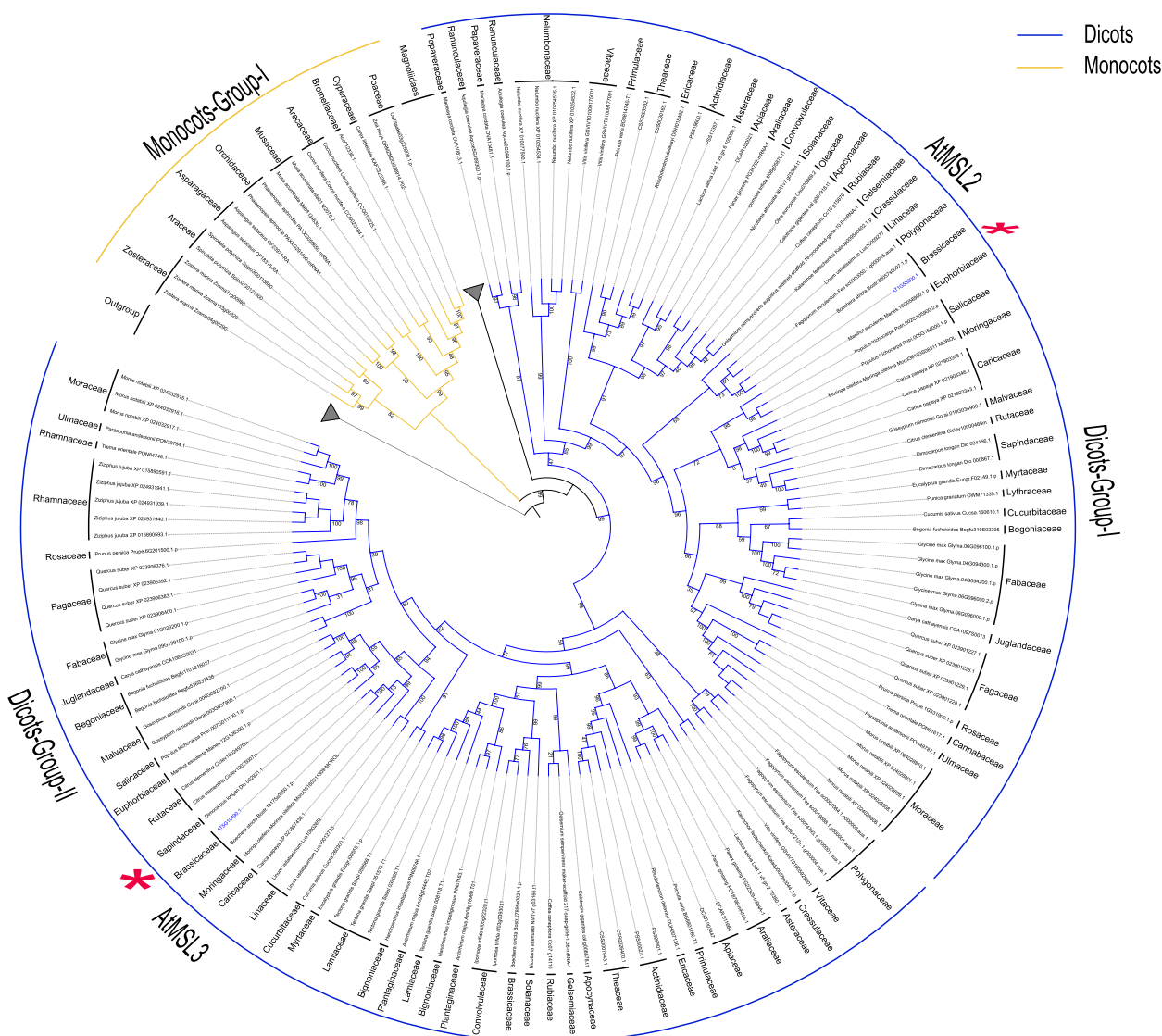
Fig. 3 A brief phylogenetic tree showing the MSL2/3 clade in angiosperms. Only selected species were included to represent each order. The topology shows that MSL2/3 in dicots can be clearly classified into two clades, Dicot-Group-I and Dicot-Group-II. Monocots, yellow; Dicots, blue. The outgroup and magnoliids collapsed into a grey triangle

species. Plant MSLs were divided into three clades before the emergence of green plants and after the plant-bacteria/protozoa/fungi split (Figure S1). This phylogenetic divergence consistent with its distinct subcellular localizations. Several previous studies classified plant MSLs into two classes with limited plant species [37, 38]. In this study, we performed a more comprehensive analysis with more plant species and displayed a more accurate phylogenetic relationship of plant MSLs (Figs. 1, 2 and 3). Like in higher plants, three MSL proteins were identified and divided into three different clades in chlorophyte, suggesting that MSLs may have gone through diversification to have sophisticated localizations and

functions in unicellular algae. During evolution, the MSL family expanded and formed four clades in seed plants. Based on the comprehensive analysis, we propose an evolutionary model for MSL in green plants (Fig. 8). Three ancestors of MSL exist in early hydrobiontic algae. Subsequently, these three MSL clades evolved independently during the evolution in land plants. Clade I is monophyletic and most plants have one homolog of MSL1. The expansion of the MSL proteins in clade II occurred in dicots, leading to two major branches of MSL2 and MSL3. The expansion of the MSL proteins in clade III occurred in seed plants, leading to two major branches of MSL4-8 and MSL9/10. In addition, the divergence of
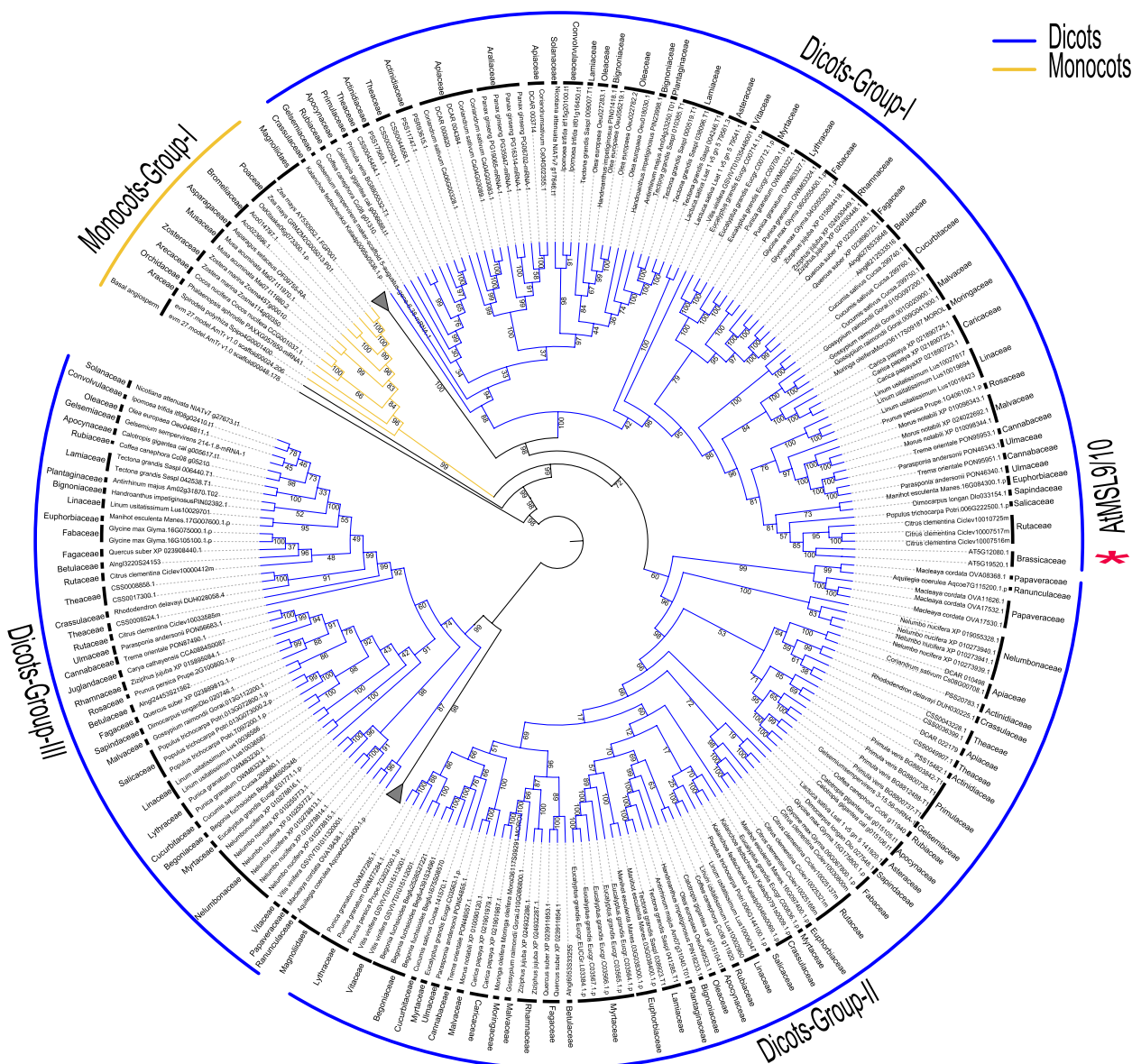
Fig. 4 Phylogenetic relationship within the MSL4-8 clade. Only selected species were included to represent each order. The topology shows that MSL4-8 in dicots can be clearly classified into two clades, Dicot-Group-I and Dicot-Group-II. Monocots, yellow; Dicots, blue. The outgroup and magnoliids collapsed into a grey triangle

MSL2/3 occurred after the monocot-dicot plant split; the divergence of MSL4/5/6/7/8 and MSL9/10 was the latest, before the emergence of Brassicaceae and after the Cleomaceae-Brassicaceae split. Within the MSL4-8 and MSL9/10 clades, a large expansion occurred in dicots, resulting in the formation of 2 and 3 subclades, respectively. This work provides insights that guide future investigations of MSL function in model and non-model organisms.

Insights into structure and functional diversification of MSLs

MSLs displayed diversity in structure, ion selectivity and physiological functions [21, 35, 38, 39, 41]. MSL proteins are characterized by a varying number of N-terminal transmembrane (TM) helices followed by a large hydrophilic cytoplasmic domain comprised primarily of ß-sheets (Fig. 1) [16, 39]. Group II MSLs have an ectoplasmic N terminus, while group III MSLs have a large cytoplasmic N terminus (Fig. 1). The key feature of the MSL proteins is the pore-lining TM helix (TM3 in EcMscS, TM5 in AtMSLs), which is broken into two parts TM3a (TM5a) and TM3b (TM5b). TM5a of AtMSL1 is rich in glycine and alanine residues, which is conserved in group I MSLs and similar to EcMscS (Fig. 6) [39]. However, multiple phenylalanine residues are rich in TM5a of group III MSLs, which is different to MSLs of group I and group II [42]. Two non-charged residues Q112 and G113 located

Zhang *et al. BMC Plant Biology*   (2023) 23:462

Page 8 of 14

Fig. 5 Phylogenetic relationship within the MSL9/10 clade. Only selected species were included to represent each order. The topology shows that MSL9/10 in dicots can be clearly classified into three clades, Dicot-Group-I, Dicot-Group-II and Dicot-Group-III. Monocots, yellow; Dicots, blue. The outgroup and magnoliids collapsed into a grey triangle

in the kink region of EcMscS and are essential for channel conductance [16]. However, differential polar residues are identified in the kink region of plant MSLs (Fig. 9). R326 and D327 in MSL1 group, R280 and E281 in MSL2/3 group, and G556 and N557 in MSL4-10 group, were identified and proved to be essential for channel conductance [39, 42–44]. Differential numbers of hydrogen bonds were formed between these differential residues, with 3, 2, 1, 1 pairs of hydrogen bonds were formed in Q-G, R-D, R-E, and G-N respectively. These residues were proposed to play important roles in modulating channel state

stabilities and transitions [39, 44]. AtMSL1 showed strong rectification compared to AtMSL8/10, whether these channel rectification differences were correlated with pore-lining residues need to be tested in the future.

In addition, many group-specific residues were identified and proved to function in MSL molecular and biological role (Fig. 9). The A324 and L329 residues are conserved in MSL1 group and work like a switch for gating and closing of the channel [39]. However, these two residues were not conserved in MSL group II and III. V273 and L277 of AtMSL2 are conserved in group II

Zhang *et al. BMC Plant Biology*   (2023) 23:462

Page 9 of 14
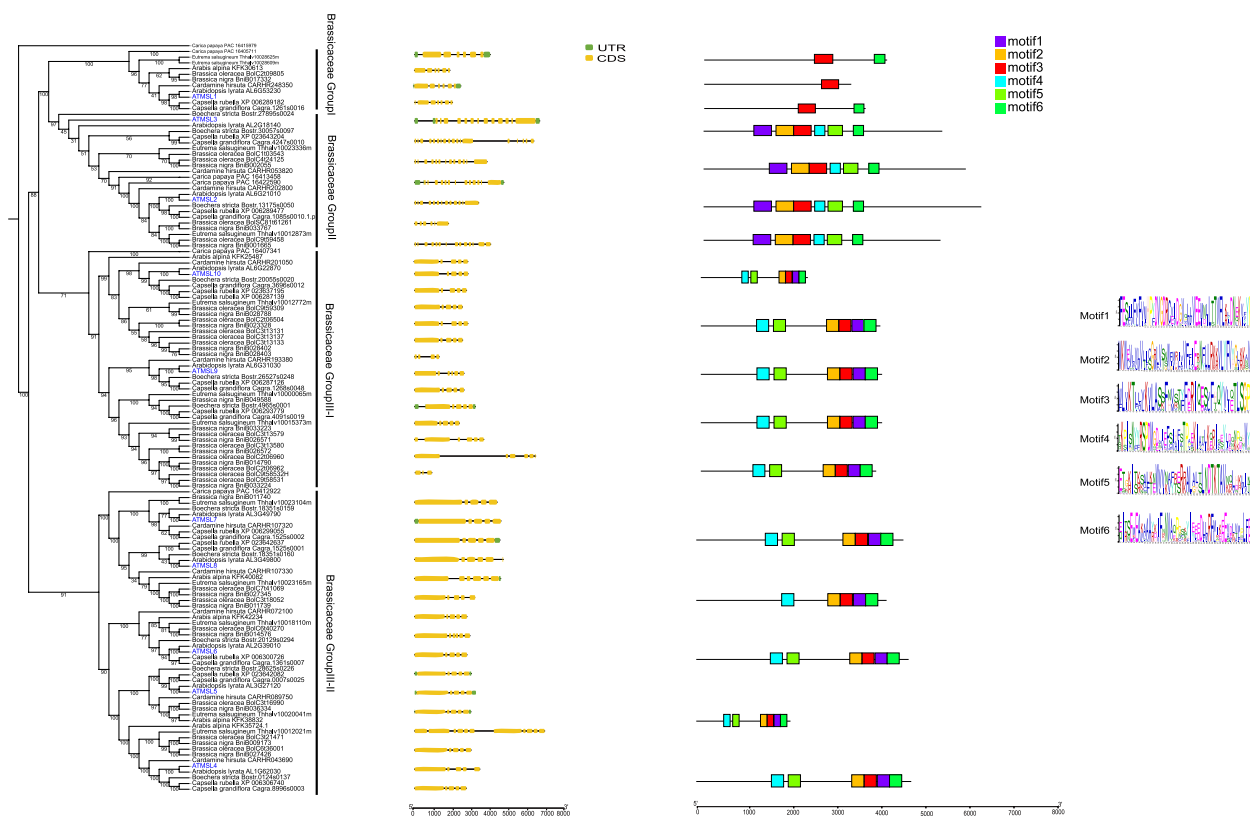
Fig. 6 Gene structure, conserved motif and phylogenetic analysis of MSLs among Brassicaceae species. The phylogenetic tree was constructed using IQ-TREE with the parameter '-m MFP -bb/alrt 1000' and 1000 ultra bootstrap replicates. The green boxes represent UTRs, yellow boxes represent CDSs and thin black lines represent introns. The motif in MSL proteins were identified by MEME program. Different motif numbered 1–6 has different colors

and are required for proper plastic localization of MSL2 [43]. F553 and I554 residues in AtMSL10 are essential for channel conductance and the stability of the open state of the channel [42]. The F553 is conserved in MSL group III and the higher plants of MSL group I, but in MSL group II it changed to leucine (Fig. 9). Seven phosphorylation sites were identified in the N-terminus of AtMSL10 and the phosphorylation regulation of these sites was critical for its function in inducing cell death [31]. These phosphorylation sites were conserved in MSL10 orthologs of Brassicaceae family, but not conserved in MSL group I and II. Therefore, though the conserved pore-lining helices of MSLs are conserved, the essential residues vary among different MSL groups, illustrating their differential channel activities and biological functions.

In plants, many *MSL* genes display different expression patterns in different organs and under different stress conditions. AtMSL2 and AtMSL3 colocalized in plastid and function redundantly [22, 26, 27]. However, *AtMSL3* displayed a higher expression than *AtMSL2*, indicating the functional divergence between *AtMSL3* and *AtMSL2*. *AtMSL8* is specially expressed in pollen [28, 29], and

AtMSL10 and AtMSL9 are highly expressed in root [23, 30]. In rice, most *MSL* genes displayed specifical expression in reproductive stages [33]. In addition, many *MSLs* displayed stress-responsive expression, indicating the important role of MSLs in stress tolerance. For example, *AtMSL2* and *AtMSL3* displayed high expression under temperature and drought stresses, and *AtMSL1* was increased under drought and waterlogging stresses [41]. In wheat, the majority of *MSL* genes were upregulated under drought, heat and heat drought stresses. Under salt stress, different *MSL* genes displayed different expression patterns, with *MSL* genes in group I were upregulated while the majority of *MSL* genes in group II were downregulated. These results indicate the different roles of *MSLs* in different tissues and under different stress conditions.

Methods
Data sources and sequence acquisition
A total of 176 species were selected and the corresponding genome and proteome sequences were obtained from public databases including NCBI (https://www.ncbi.nlm.nih.

Zhang *et al. BMC Plant Biology*   (2023) 23:462

Page 10 of 14
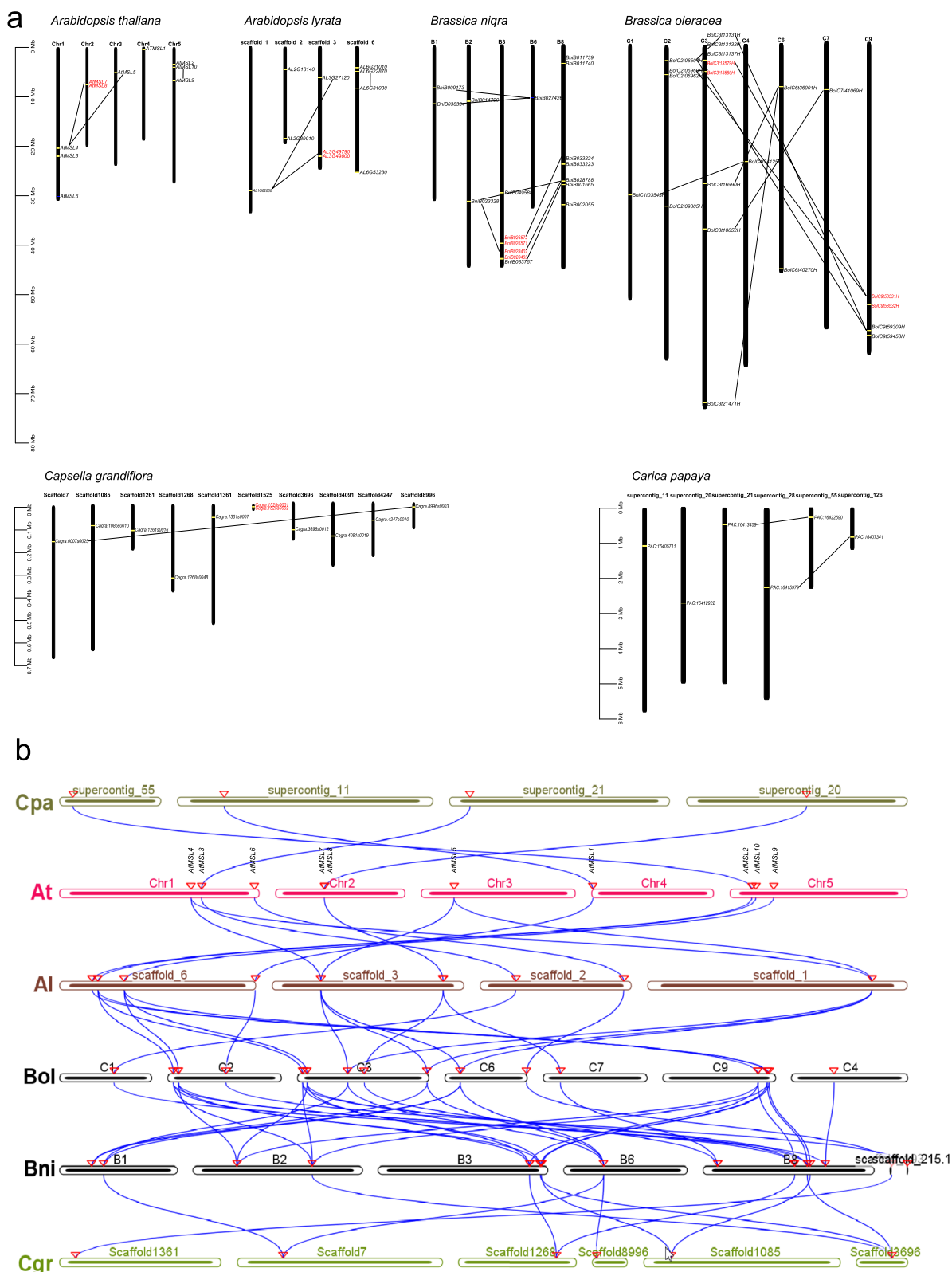
Fig. 7 Synteny analysis of Brassicaceae MSL Proteins. **a** Genomic distribution of *MSL* genes across Brassicaceae species. Collinearity genes are linked by black line and tandem genes are marked by red color. **b** Multicollinearity analysis of *MSL* genes among different Brassicales species. The blue lines represent the MSL syntenic genes among *Carica papaya* (*Cpa*), *Arabidopsis thaliana* (*At*), *Arabidopsis lyrate* (*Al*), *Brassica niqra* (*Bni*), *Brassica oleracea* (*Bol*), and *Capsella grandiflora* (*Cgr*) genomes. The location of the *MSL* gene on the chromosome is marked by a triangle

Zhang *et al. BMC Plant Biology*   (2023) 23:462

Page 11 of 14
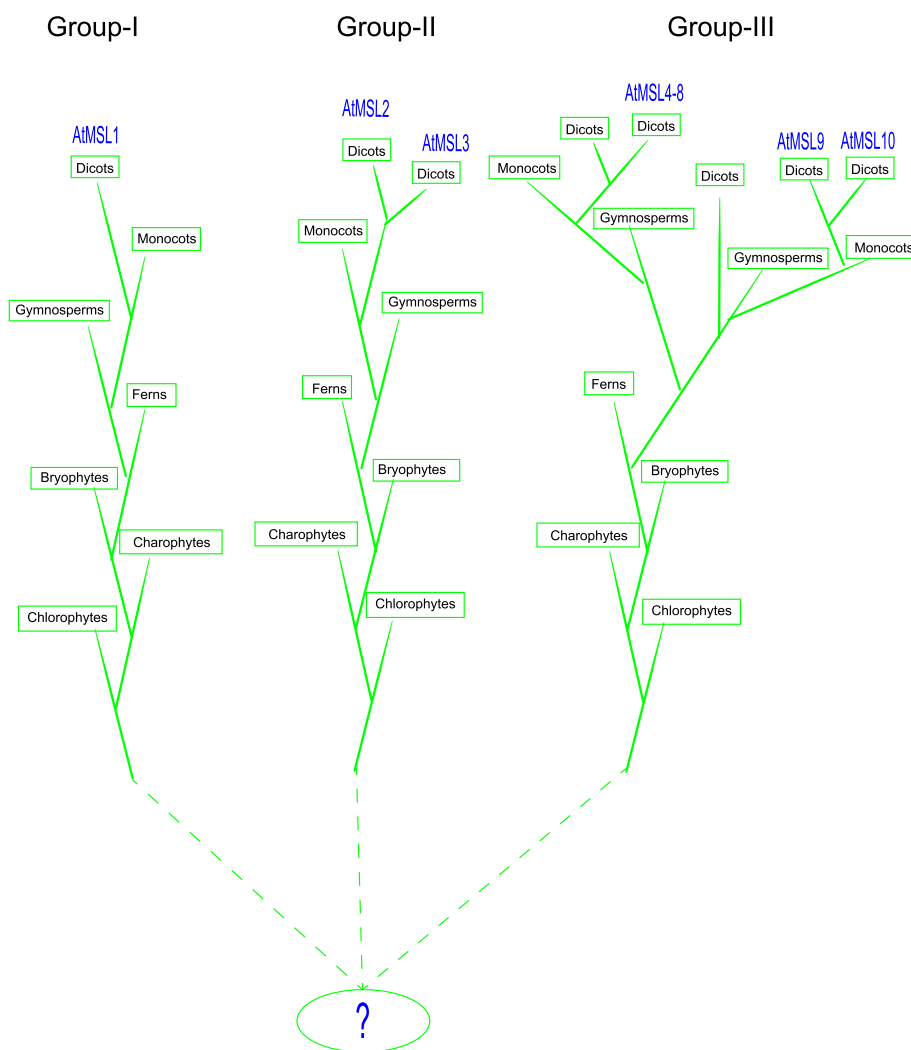
Fig. 8 A proposed evolutionary model of MSLs in plants. The model is based on the phylogeny of MSLs and the cladogram of green plant species. The origin of plant MSLs was traced back to chlorophytes. During the evolution, MSLs were divided into four major groups in angiosperms

gov/), Phytozome v13. 0 (https://phytozome.jgi.doe.gov) and Gigadb (www.gigadb.org) (Supplementary Table 1).

Two methods were used to identify *MSL* genes in plants. First, HMMER search (E-value = 1e − 10) was employed with the Hidden Markov Model profile of MscS channel domain (PF00924) to search the local Databases. Second, the amino acid sequences of *Escherichia coli* MscS and *Arabidopsis thaliana* MSL members were used to run a Basic Local Alignment Search Tool algorithms (BLASTP) search against the protein database with an E-value less than 10^{-6}. The putative MSLs were further validated with online tools CDD (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi/) [45], HMM (https://hmmer.org/) [46] and SMART (https://smart.embl-heidelberg.de/) [47]. The transmembrane domain

(TM) of MSLs was predicted using TMHMM Server v 2.0 (http://www.cbs.dtu.dk/ services/TMHMM-2.0/). Only sequences with PF00924 and TM domains were retained. In total, 2113 MSL proteins were identified and used for further analysis (Table S1).
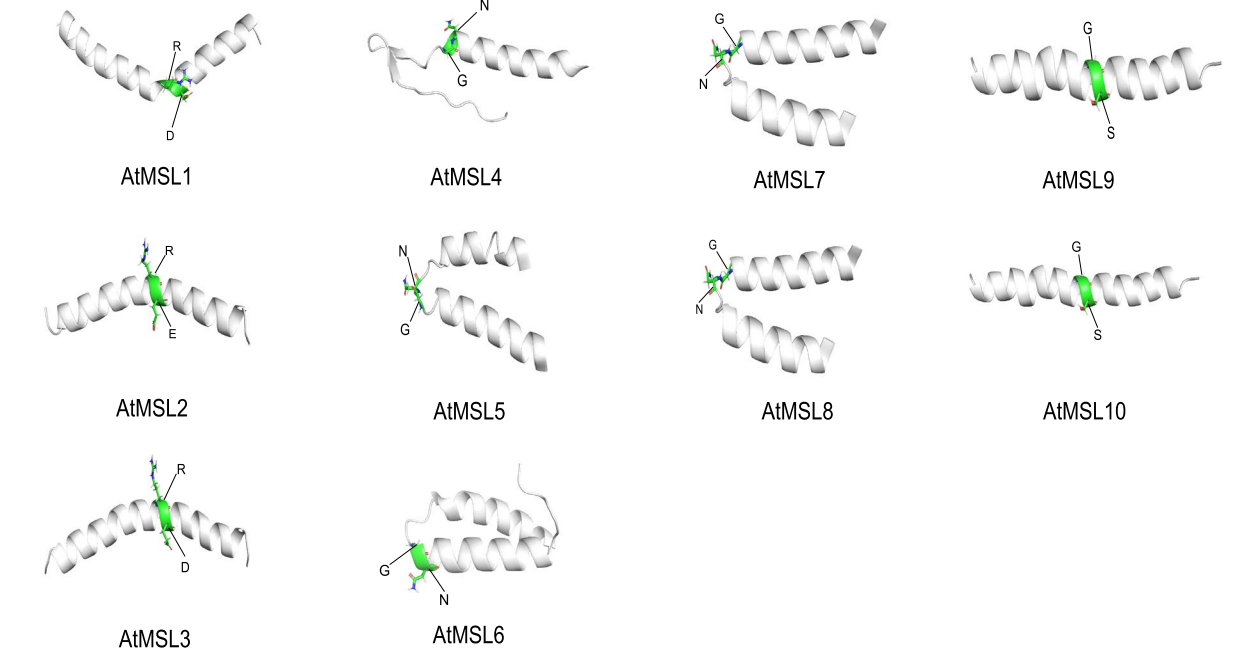
Multiple sequence alignment, protein structure predictions and phylogenetic analysis

Multiple sequence alignments (MSA) of MSLs were performed with MAFFT software [48]. The putative structures of AtMSL2-10 were predicted with I-TASSER prediction server and AlphaFold using the AtMSL1 crystal structure as a template [39, 49]. The 3D models were validated using ProSA [50]. The crystal structure was visualized with open software PyMod [51].

Zhang *et al. BMC Plant Biology*   (2023) 23:462

Page 12 of 14
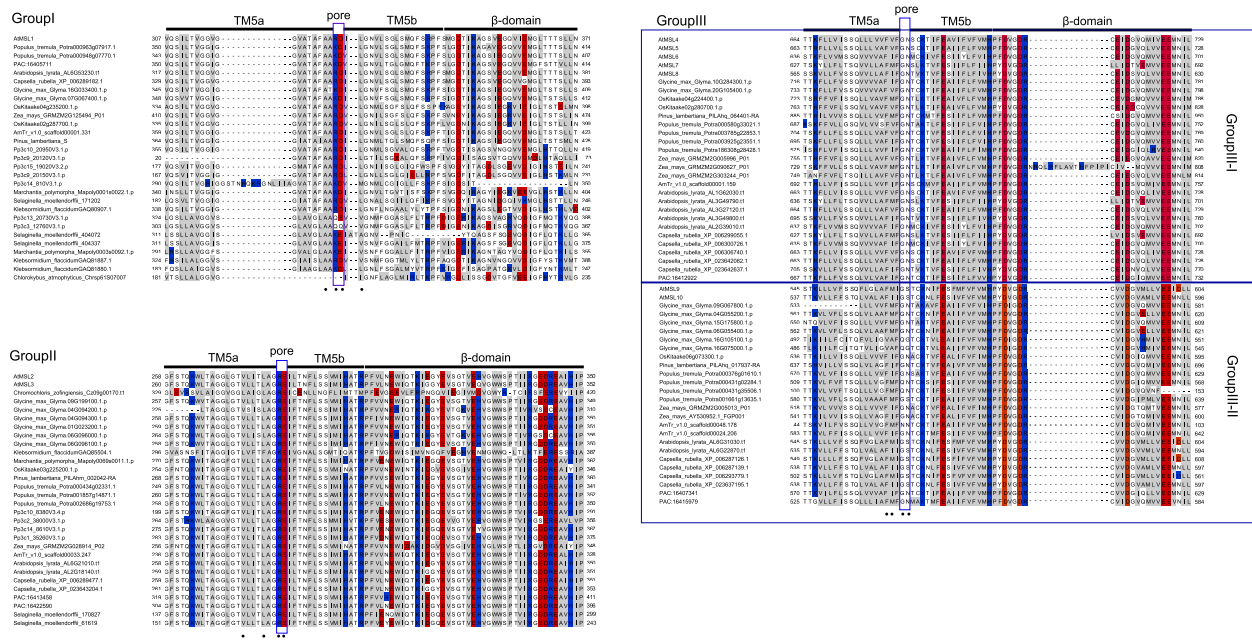
Fig. 9 Multiple sequence alignment and conserved motifs in MSL proteins. **a** Alignment of pore-lining helices from MSL proteins. Sequences corresponding to TM3a, TM3b, and the ß-domain of *E.Coli* MscS are indicated by lines. Polar residues are white, non-polar residues gray, positively charged residues blue, negatively charged residues red. **b** Side view of the predicted pore-lining domain of the AtMSL1-10 created with I-TASSER. The side chains that predicted to form a kink are indicated with green sphere

To explore the evolutionary origin of MSL proteins in green plants and to elucidate the phylogenetic relationship of MSLs in angiosperms, we constructed 3 phylogenetic trees based on taxonomy: Dataset I contains 7 non-angiosperms and 9 angiosperms, Dataset II contains 156 angiosperms, Dataset III contains 11 Brassicales species (Table S3). The phylogenetic trees were constructed based on the core amino acid regions corresponding to

Zhang *et al. BMC Plant Biology*   (2023) 23:462

Page 13 of 14

transmembrane domain 3 and the adjacent consensus sequence of MSLs. The Bayesian trees were constructed with MrBayes 3.2.1 using the fixed Whelan and Goldman model with six Markov chains until the average standard deviation of split frequencies was < 0.05 (Dataset I tree: 600,000 generations, Dataset II tree: 700,000 generations and Dataset IIII tree: 500,000 generations) [52]. The maximum likelihood (ML) phylogenetic tree was constructed with IQ-TREE with the parameter '-m MFP -bb/alrt 1000' and 1000 ultra bootstrap replicates [53].

Subcellular localization prediction, conserved motif and gene structure analysis

The subcellular localization of MSLs was predicted with Wolf PSORT (https://wolfpsort.hgc.jp). The exon–intron structures were displayed through Gene Structure Display Server (GSDS) (http://gsds.cbi.pku.edu.cn/index.php). Conserved motifs were predicted with MEME (http://meme.nbcr.net/meme3/mme.html) [54]. 6 motifs with a minimum and maximum length of 50 and 200 have been taken into account.

Synteny analysis

Homolog pairs between species and within a certain species were identified using the all-to-all BLASTP method, and syntenic blocks were inferred using MCScanX with the default parameters [55]: E-value, 1e-10; BLAST hists, 5. The syntonic map was displayed using CIRCOS with the putative duplicated genes were linked by the connection lines.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-023-04479-2.

Additional file 1: Supplementary Figure 1. Phylogenetic analysis of MSL proteins among fungi, protozoa, bacteria and plants.

Additional file 2: Supplementary Figure 2. Phylogenetic relationship within the Group III of MSL proteins.

Additional file 3: Supplementary Table 1. Green plants, sources of sequence and genome version information used in this study.

Additional file 4: Supplementary Table 2. The number of MSL proteins in the identified plants.

Additional file 5: Supplementary Table 3. The number of exons/introns in Brassicaceae MSL proteins.

Additional file 6: Supplementary Table 4. The detailed information of MSL proteins identified in the present study.

Authors' contributions
The study was conceived and directed by Zaibao Zhang. Zaibao Zhang, and Fan Ye wrote the manuscript. Fan Ye, Tao Xiong, Jiahui Chen, Jiajia Cao, and Yurui Chen performed the identification of *MSL* genes, protein structure and evolution analysis. Sushuang Liu participated in manuscript preparation and revision. All the authors read and approved the final manuscript. All authors consent for publication.

Availability of data and materials
The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request. All raw sequencing data were downloaded from public database. The detailed information could be found in Supplementary Table 1.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Cox CD, Nakayama Y, Nomura T, Martinac B. The evolutionary 'tinkering' of MscS-like channels: generation of structural and functional diversity. Pflugers Arch. 2015;467(1):3–13.
2. Anishkin A, Kung C. Microbial mechanosensation. Curr Opin Neurobiol. 2005;15(4):397–405.
3. Malcolm HR, Maurer JA. The mechanosensitive channel of small conductance (MscS) superfamily: not just mechanosensitive channels anymore. ChemBioChem. 2012;13(14):2037–43.
4. Pivetti CD, Yen MR, Miller S, Busch W, Tseng YH, Booth IR, Saier MH Jr. Two families of mechanosensitive channel proteins. Microbiol Mol Biol Rev. 2003;67(1):66–85 table of contents.
5. Martinac B. Mechanosensitive channels in prokaryotes. Cell Physiol Biochem. 2001;11(2):61–76.
6. Honoré E. The neuronal background K2P channels: focus on TREK1. Nat Rev Neurosci. 2007;8(4):251–61.
7. Kurusu T, Yamanaka T, Nakano M, Takiguchi A, Ogasawara Y, Hayashi T, Iida K, Hanamata S, Shinozaki K, Iida H, et al. Involvement of the putative Ca^{2+}-permeable mechanosensitive channels, NtMCA1 and NtMCA2, in Ca^{2+} uptake, Ca^{2+}-dependent cell proliferation and mechanical stress-induced gene expression in tobacco (Nicotiana tabacum) BY-2 cells. J Plant Res. 2012;125(4):555–68.
8. Volkers L, Mechioukhi Y, Coste B. Piezo channels: from structure to function. Pflugers Arch. 2015;467(1):95–9.
9. Booth IR, Rasmussen T, Edwards MD, Black S, Rasmussen A, Bartlett W, Miller S. Sensing bilayer tension: bacterial mechanosensitive channels and their gating mechanisms. Biochem Soc Trans. 2011;39(3):733–40.
10. Kung C, Martinac B, Sukharev S. Mechanosensitive channels in microbes. Annu Rev Microbiol. 2010;64:313–29.
11. Schumann U, Edwards MD, Rasmussen T, Bartlett W, van West P, Booth IR. YbdG in Escherichia coli is a threshold-setting mechanosensitive channel with MscM activity. Proc Natl Acad Sci USA. 2010;107(28):12664–9.
12. Edwards MD, Black S, Rasmussen T, Rasmussen A, Stokes NR, Stephen TL, Miller S, Booth IR. Characterization of three novel mechanosensitive channel activities in Escherichia coli. Channels (Austin). 2012;6(4):272–81.
13. Cox CD, Bavi N, Martinac B. Bacterial Mechanosensors. Annu Rev Physiol. 2018;80:71–93.
14. Levina N, Tötemeyer S, Stokes NR, Louis P, Jones MA, Booth IR. Protection of Escherichia coli cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity. EMBO J. 1999;18(7):1730–7.

Zhang *et al. BMC Plant Biology*   (2023) 23:462

Page 14 of 14

15. Boer M, Anishkin A, Sukharev S. Adaptive MscS gating in the osmotic permeability response in E. coli: the question of time. Biochemistry. 2011;50(19):4087–96.

16. Bass RB, Strop P, Barclay M, Rees DC. Crystal structure of Escherichia coli MscS, a voltage-modulated and mechanosensitive channel. Science. 2002;298(5598):1582–7.

17. Balleza D, Gómez-Lagunas F. Conserved motifs in mechanosensitive channels MscL and MscS. Eur Biophys J. 2009;38(7):1013–27.

18. Nomura T, Sokabe M, Yoshimura K. Lipid-protein interaction of the MscS mechanosensitive channel examined by scanning mutagenesis. Biophys J. 2006;91(8):2874–81.

19. Malcolm HR, Heo YY, Elmore DE, Maurer JA. Defining the role of the tension sensor in the mechanosensitive channel of small conductance. Biophys J. 2011;101(2):345–52.

20. Haswell ES. MscS-Like Proteins in Plants. Curr Top Membr. 2007;58(06): 329–59.

21. Hamilton ES, Schlegel AM, Haswell ES. United in diversity: mechanosensitive ion channels in plants. Annu Rev Plant Biol. 2015;66:113–37.

22. Haswell ES, Meyerowitz EM. MscS-like proteins control plastid size and shape in Arabidopsis thaliana. Current biology : CB. 2006;16(1):1–11.

23. Haswell ES, Peyronnet R, Barbier-Brygoo H, Meyerowitz EM, Frachisse JM. Two MscS homologs provide mechanosensitive channel activities in the Arabidopsis root. Curr Biol. 2008;18(10):730–4.

24. Guerringue Y, Thomine S, Frachisse JM. Sensing and transducing forces in plants with MSL10 and DEK1 mechanosensors. FEBS Lett. 2018;592(12):1968–79.

25. Lee CP, Maksaev G, Jensen GS, Murcha MW, Wilson ME, Fricker M, Hell R, Haswell ES, Millar AH, Sweetlove LJ. MSL1 is a mechanosensitive ion channel that dissipates mitochondrial membrane potential and maintains redox homeostasis in mitochondria during abiotic stress. Plant J. 2016;88(5):809–25.

26. Veley KM, Marshburn S, Clure CE, Haswell ES. Mechanosensitive channels protect plastids from hypoosmotic stress during normal plant growth. Curr Biol. 2012;22(5):408–13.

27. Wilson ME, Jensen GS, Haswell ES. Two mechanosensitive channel homologs influence division ring placement in Arabidopsis chloroplasts. Plant Cell. 2011;23(8):2939–49.

28. Hamilton ES, Haswell ES. The tension-sensitive ion transport activity of MSL8 is critical for its function in pollen hydration and germination. Plant Cell Physiol. 2017;58(7):1222–37.

29. Hamilton ES, Jensen GS, Maksaev G, Katims A, Sherp AM, Haswell ES. Mechanosensitive channel MSL8 regulates osmotic forces during pollen hydration and germination. Science. 2015;350(6259):438–41.

30. Peyronnet R, Haswell ES, Barbier-Brygoo H, Frachisse JM. AtMSL9 and AtMSL10: Sensors of plasma membrane tension in Arabidopsis roots. Plant Signal Behav. 2008;3(9):726–9.

31. Basu D, Shoots JM, Haswell ES. Interactions between the N- and C-termini of the mechanosensitive ion channel AtMSL10 are consistent with a three-step mechanism for activation. J Exp Bot. 2020;71(14):4020–32.

32. Zou Y, Chintamanani S, He P, Fukushige H, Yu L, Shao M, Zhu L, Hildebrand DF, Tang X, Zhou JM. A gain-of-function mutation in Msl10 triggers cell death and wound-induced hyperaccumulation of jasmonic acid in Arabidopsis. J Integr Plant Biol. 2016;58(6):600–9.

33. Saddhe AA, Kumar K. In silico identification and expression analysis of MscS like gene family in rice. Plant Gene. 2015;1:8–17.

34. Silvia GHuldS, Adriana PdS, Tania MI, Eduardo GdS, Talita Cu. Genome-wide analysis of mechanosensitive channel of small conductance (MscS)-like gene family in common bean. Afr J Biotechnol. 2016;15(15):580–92.

35. Hamant O, Haswell ES. Life behind the wall: sensing mechanical cues in plants. BMC Biol. 2017;15(1):59.

36. Basumatary N, Kumari A, Sarmah J. Genome-wide identification of MscS like gene family in Cicer arietinum using bioinformatics approach. Ann Plant Sci. 2018;7(2):2002.

37. Kaur A, Taneja M, Tyagi S, Sharma A, Singh K, Upadhyay SK. Genome-wide characterization and expression analysis suggested diverse functions of the mechanosensitive channel of small conductance-like (MSL) genes in cereal crops. Sci Rep. 2020;10(1):16583.

38. Kaur A, Madhu, Upadhyay SK. Mechanosensitive ion channels in plants. In: Calcium transport elements in plants. 2021. p. 267–79.

39. Li Y, Hu Y, Wang J, Liu X, Zhang W, Sun L. Structural insights into a plant mechanosensitive ion channel MSL1. Cell Rep. 2020;30(13):4518-4527.e4513.

40. Moore RC, Purugganan MD. The early stages of duplicate gene evolution. Proc Natl Acad Sci U S A. 2003;100(26):15682–7.

41. Jiang W, Tong T, Chen X, Deng F, Zeng F, Pan R, Zhang W, Chen G, Chen ZH. Molecular response and evolution of plant anion transport systems to abiotic stress. Plant Mol Biol. 2022;110(4–5):397–412.

42. Maksaev G, Shoots JM, Ohri S, Haswell ES. Nonpolar residues in the presumptive pore-lining helix of mechanosensitive channel MSL10 influence channel behavior and establish a nonconducting function. Plant Direct. 2018;2(6):e00059.

43. Jensen GS, Haswell ES. Functional analysis of conserved motifs in the mechanosensitive channel homolog MscS-Like2 from Arabidopsis thaliana. PLoS One. 2012;7(6):e40336.

44. Schlegel AM, Haswell ES. Charged pore-lining residues are required for normal channel kinetics in the eukaryotic mechanosensitive ion channel MSL1. Channels (Austin). 2020;14(1):310–25.

45. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, et al. CDD: NCBI's conserved domain database. Nucleic Acids Res. 2015;43(Database issue):D222-226.

46. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 2016;44(D1):D279-285.

47. Letunic I, Doerks T, Bork P. SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Res. 2012;40(Database issue):D302-305.

48. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.

49. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc. 2010;5(4):725–38.

50. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007;35(Web Server issue):W407-410.

51. Bramucci E, Paiardini A, Bossa F, Pascarella S. PyMod: sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL. BMC Bioinformatics. 2012;13 Suppl 4(Suppl 4):S2.

52. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003;19(12):1572–4.

53. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQTREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020;37:1530–34. https://doi.org/10.1093/molbev/msaa015.

54. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37(Web Server issue):W202-208.

55. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49.

Publisher's Note