



Molecular Evolution of GDP-D-Mannose Epimerase (*GME*), a Key Gene in Plant Ascorbic Acid Biosynthesis

Junjie Tao^{1,2}, Han Wu^{1,2}, Zhangyun Li^{1,2}, Chunhui Huang^{1,2} and Xiaobiao Xu^{1,2*}

¹ College of Agronomy, Jiangxi Agricultural University, Nanchang, China, ² Institute of Kiwifruit, Jiangxi Agricultural University, Nanchang, China

OPEN ACCESS

Edited by:

Juan Caballero,
Universidad Autónoma de Querétaro,
Mexico

Reviewed by:

Cheong Xin Chan,
The University of Queensland,
Australia
Hao Wang,
University of Georgia, United States

*Correspondence:

Xiaobiao Xu
xiaobiaoxu@hotmail.com

Specialty section:

This article was submitted to
Bioinformatics and Computational
Biology,
a section of the journal
Frontiers in Plant Science

Received: 16 April 2018

Accepted: 17 August 2018

Published: 04 September 2018

Citation:

Tao J, Wu H, Li Z, Huang C and Xu X
(2018) Molecular Evolution
of GDP-D-Mannose Epimerase
(*GME*), a Key Gene in Plant Ascorbic
Acid Biosynthesis.
Front. Plant Sci. 9:1293.
doi: 10.3389/fpls.2018.01293

The widespread ascorbic acid (AsA) plays a vital role in plant development and abiotic stress tolerance, but AsA concentration varies greatly among different plants. GDP-D-mannose epimerase (*GME*), which catalyzes GDP-D-mannose to GDP-L-galactose or GDP-L-gulose, is a key enzyme in plant AsA biosynthesis pathway. Functions and expression patterns of *GME* have been well studied in previous works, however, little information is known about the evolutionary patterns of the gene. In this study, *GME* gene structure, corresponding conserved protein motifs and evolutionary relationships were systematically analyzed. A total of 111 *GME* gene sequences were retrieved from 59 plant genomes, which representing almost all the major lineages of Viridiplantae: dicotyledons, monocotyledons, gymnosperms, pteridophytes, bryophytes, and chlorophytes. Results showed that homologs of *GME* were widely present in Viridiplantae. *GME* gene structures were conservative in higher plants, while varied greatly in the basal subgroups of the phylogeny including lycophytes, bryophytes, and chlorophytes, suggesting *GME* gene structure might have undergone severe differentiation at lower plant and then gradually fixed as plant evolution. The basic motifs of *GME* were strongly conserved throughout Viridiplantae, suggesting the conserved function of the protein. Molecular evolution analysis showed that strong purifying selection was the predominant force in the evolution of *GME*. A few branches and sites under episodic diversifying selection were identified and most of the branches located in the subgroup of chlorophytes, indicating episodic diversifying selection at a few branches and sites may play a role in the evolution of *GME* and diversifying selection may have occurred at the early stage of Viridiplantae. Our results provide novel insights into functional conservation and the evolution of *GME*.

Keywords: ascorbic acid, *GME*, molecular evolution, L-galactose pathway, Viridiplantae

INTRODUCTION

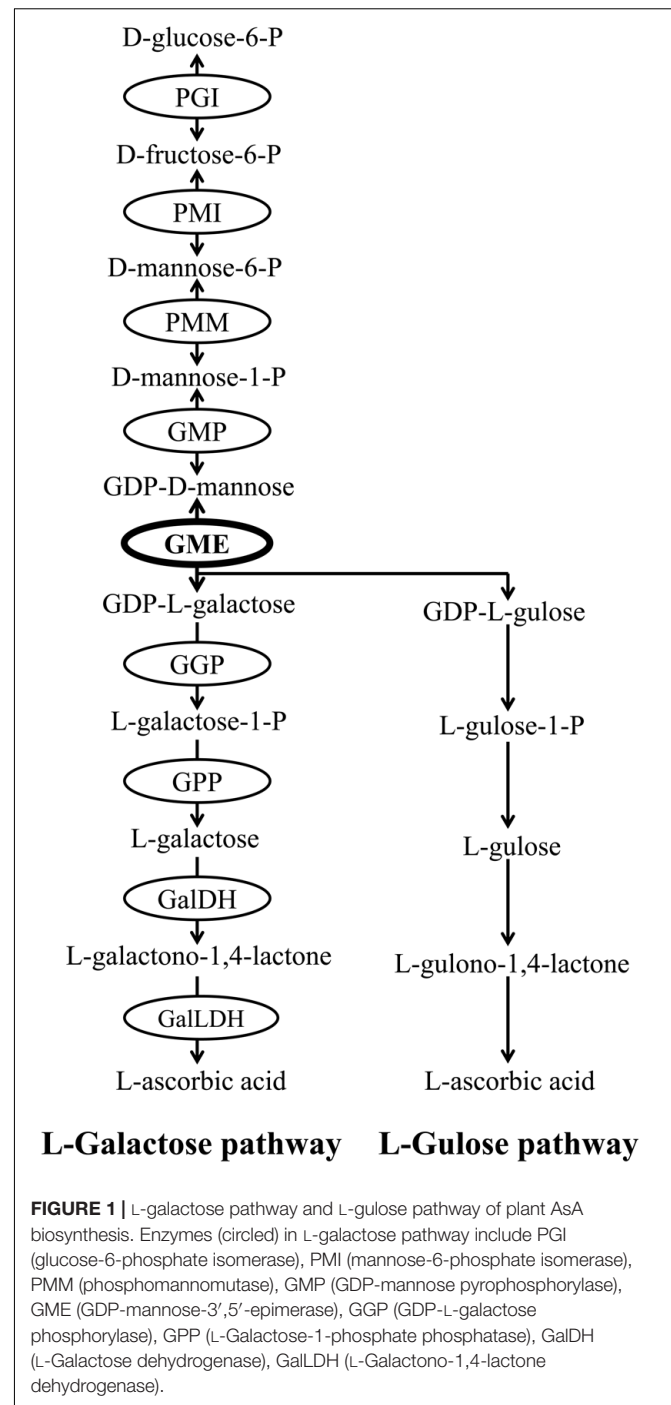
L-ascorbic acid (AsA), also well known as vitamin C or ascorbate, is an essential water-soluble micronutrient for both animals and plants. As a general antioxidant and a cofactor for enzymes, AsA plays important roles in the process of sustaining biological activities in living organisms. Earlier studies revealed that AsA is a versatile reducing agent with anticancer properties

(Ma et al., 2014; Yun et al., 2015; Doskey et al., 2016). In plant tissues, AsA is a multifunctional metabolite and has been implicated in numerous cellular processes of growth and development, including photosynthesis (Talla et al., 2011), seed set and germination (Liu et al., 2010; Ye et al., 2012), flowering (Barth et al., 2006; Kotchoni et al., 2009), senescence (Wang et al., 2011; Zhang et al., 2013) and so on. Moreover, AsA also plays a role in promoting plant tolerance to various environmental stresses (Gallie, 2012; Akram et al., 2017).

Although AsA distribute widely over Viridiplantae, different organisms have varying amounts of AsA concentration. In lower photosynthetic organism of cyanobacteria, AsA level is about 250-fold lower than that in higher plants (Tel-or et al., 1986; Obinger et al., 1998). In green algae, *Ulva compressa* and *Ulva fasciata* have very similar levels of AsA and the concentration is about $0.5 \mu\text{mol g}^{-1}$ fresh weight (FW) (Shiu and Lee, 2005; Mellado et al., 2012). In mosses, AsA concentrations are relatively closer with algae, such as the levels of AsA in *Hypnum plumaforme*, *Brachythecium velutinum*, and *Marchantia polymorpha* are $0.1\text{--}0.6 \mu\text{mol g}^{-1}$ FW, $0.25\text{--}0.5 \mu\text{mol g}^{-1}$ FW, and $0.3 \mu\text{mol g}^{-1}$ FW, respectively (Paciolla and Tommasi, 2003; Sun et al., 2010). However, the ascorbate concentrations are still about several-fold lower in algae and mosses than in higher plants, which generally have a wide range of variability of AsA contents from 2 to $138 \mu\text{mol g}^{-1}$ FW (Gest et al., 2013). Furthermore, the amount of ascorbate has been shown to have considerable variation even within a genus or species. For example, in the genus *Actinidia*, *A. chinensis* var. *chinensis* has concentrations about $2.84\text{--}23.86 \mu\text{mol g}^{-1}$ FW, and *A. chinensis* var. *deliciosa* contains between 2.84 and $14.2 \mu\text{mol g}^{-1}$ FW, while *A. eriantha* can have concentrations over $28.4 \mu\text{mol g}^{-1}$ FW (Huang et al., 2004). The AsA concentrations in different plants are largely affected by external environmental factors, such as light intensity, altitude, and temperature (Gest et al., 2013). In addition, internal molecular mechanisms, such as biosynthesis, oxidation and recycling, also play important roles in maintaining optimal levels of ascorbate in plant growth and development.

Ascorbic acid biosynthetic pathways have been well documented in plants. At least four biosynthetic pathways have been identified in plants, including L-galactose pathway (Wheeler et al., 1998), L-gulose pathway (Wolucka and Van Montagu, 2003), D-Galacturonate pathway (Agius et al., 2003) and myo-Inositol pathway (Lorence et al., 2004). Among these four pathways, the L-galactose pathway has been fully elucidated and is regarded as the primary route of ascorbate biosynthesis in most plant species (Bulley and Laing, 2016). Nine structural genes are involved in L-galactose pathway and all of them have been identified and cloned from various plants (Figure 1).

GDP-D-mannose 3',5'-epimerase (GME) is considered to be a key regulator of ascorbate accumulation and cell wall biosynthesis. Plant GME release two epimerization products from GDP-D-mannose, one is GDP-L-galactose in L-galactose pathway and another is GDP-L-gulose in an alternative pathway of L-gulose (Figure 1; Wheeler et al., 1998; Wolucka and Van Montagu, 2003). GME is the most conserved protein in ascorbate biosynthesis pathway (Wolucka and Montagu, 2007), and its



function has been demonstrated in many higher plants. For example, the AsA content is decreased significantly and plant growth and development is impaired severely in GME-silenced tomato (Gilbert et al., 2009; Voxeur et al., 2011). Meanwhile, AsA accumulation is increased and abiotic stresses tolerance is enhanced obviously in GME-overexpression tomato (Zhang et al., 2011). Furthermore, GME also regulates plant vegetative growth and reproductive development in *Arabidopsis* (Qi et al., 2017). However, earlier studies in some plants, such as kiwifruit

(Bulley et al., 2009; Li et al., 2013) and peach (Imai et al., 2009), showed that GME alone may not determine AsA levels. As a result, further studies from more-broadly sampled taxa are needed to improve our understandings of the exact function of GME in regulating plant AsA biosynthesis.

Previous studies of *GME* mainly focused on gene expression patterns, regulatory mechanisms and functional analysis, while the evolutionary patterns of *GME* in plants are still unclear till now. In the present study, *GME* sequences were retrieved from 59 plant species which covering the main lineages of Viridiplantae and the molecular evolution of *GME* was investigated by phylogenetic analyses and molecular evolutionary approaches. Our results provide novel insights into the evolution and the function of *GME*, particularly in the context of AsA biosynthesis.

MATERIALS AND METHODS

Sequence Acquisition and Characterization

The genomic DNA sequences, coding sequences and corresponding protein sequences of plant *GME* genes used in this study were retrieved from the public databases Phytozome v12.1¹ and National Center of Biotechnology Information (NCBI)². To identify genes encoding GME proteins in Viridiplantae, the AtGME (AT5G28840) amino acid sequence of *Arabidopsis thaliana* and AeGME (MG383560) protein sequences of *A. eriantha* were used as queries to perform BLASTP searches with default algorithm parameters in Phytozome v12.1 database (Goodstein et al., 2012). Additionally, GME protein sequences of *A. deliciosa* and *A. rufa*, which were obtained from NCBI using AeGME sequence as a query, were also used in following analyses. The identical, redundant, and incomplete sequences were removed manually from the downloaded data set using BioEdit v7.1.3 (Hall, 1999), and only the full-length coding sequences were retained. Then, the potential plant GME protein sequences retrieved from Phytozome v12.1 and NCBI were checked and analyzed using HMMER website³ (Finn et al., 2015) and the CD-search tool in the Conserved Domain Database of NCBI⁴ (Marchler-Bauer et al., 2015) to confirm the presence of NAD-binding domain, which is a well conserved functional domain of plant GMEs (Watanabe et al., 2006; Siow et al., 2013).

Multiple Sequence Alignment

Multiple sequence alignment for all the selected plant GME amino acid sequences was firstly generated using MAFFT v7.158 (Katoh and Standley, 2013) by the auto option, after which they were manually adjusted using BioEdit v7.1.3 (Hall, 1999). Then, the alignment of the amino acid sequences was converted into the alignment of nucleotide sequences in fixed codon positions utilizing PAL2NAL⁵ (Suyama et al., 2006). Finally, Gblocks v0.91b

(Castresana, 2000) was used to identify conserved blocks and regions that had gaps in more than 50% of the sequences in the multiple alignment of nucleotide sequence were removed.

Recombination Detection

Phylogenetic reconstruction and molecular evolutionary analysis can be profoundly influenced by recombination (Posada and Crandall, 2002; Anisimova et al., 2003; Shriener et al., 2003). To avoid the impact of recombination, Genetic Algorithm Recombination Detection (GARD) (Kosakovsky Pond et al., 2006) method implemented in Datamonkey website⁶ (Delpont et al., 2010) was used to test whether there was any evidence of recombination in plant *GME* nucleotide sequences.

Phylogenetic Tree Reconstruction

The phylogenetic tree reconstruction was performed using Bayesian inference method implemented in MrBayes v3.1.2 software (Ronquist and Huelsenbeck, 2003). Prior to the phylogenetic analysis, the most suitable nucleotide substitution model was firstly determined using the AIC criterion in MrModeltest v2.3 (Nylander, 2008). The Bayesian phylogenetic estimation was conducted by running four independent Markov Chain Monte Carlo (MCMC) algorithms (three heated chains and one cold chain) for 10,000,000 generations. Trees were sampled every 100 generations and the first 25% of the samples were discarded as burn-in. Finally, the resulting phylogenetic tree was edited and visualized in iTOL⁷.

Gene Structure and Conserved Motifs Analyses

To analysis gene structure of *GME*, the Gene Structure Display Server v2.0 (GSDS v2.0⁸) was used to display the exon-intron organizations of *GME* by uploading *GME* genomic DNA sequences and their corresponding cDNA sequences to the website. Conserved motifs of GME were identified using the Multiple Expectation Maximization for Motif Elicitation (MEME) v4.12.0⁹ with the maximum number of motifs = 8 (optimized based on an iterative analysis), motif width between 15 and 50 aa, and other parameters at default setting.

Molecular Evolution Analyses

The ratio of non-synonymous (d_N) over synonymous substitutions (d_S) ($\omega = d_N/d_S$) is an effective indicator to evaluate selective pressures acting on protein-coding genes, where $\omega > 1$, $\omega < 1$, and $\omega = 1$ indicates diversifying positive selection, negative selection and neutral evolution, respectively. To investigate the influence of selective pressures acting on *GME* genes in Viridiplantae, codon-based maximum likelihood methods implemented in codeml program from PAML package version 4.7 were utilized (Yang, 2007). The reconstructed Bayesian phylogenetic tree and the alignment of codon-based nucleotide sequences were input into codeml program to

¹<https://phytozome.jgi.doe.gov/pz/portal.html>

²<https://www.ncbi.nlm.nih.gov/>

³<https://www.ebi.ac.uk/Tools/hmmer/>

⁴<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>

⁵<http://www.bork.embl.de/pal2nal/>

⁶<http://www.datamonkey.org/>

⁷<https://itol.embl.de/>

⁸<http://gsds.cbi.pku.edu.cn/>

⁹<http://alternate.meme-suite.org/tools/meme>

perform molecular evolution analyses. The site-specific models were firstly employed to examine the selective pressure at amino acid sites and identify potential sites under diversifying selection (Yang et al., 2000). Three pairs of log-likelihood ratio tests (LRTs) of M0 (one-ratio model) vs. M3 (discrete model), M1a (neutral model) vs. M2a (selection model), M7 (β) vs. M8 (β and ω) were employed to detect the variation of ω among sites. For one LRT, the difference in the number of parameters between nested models was used as degree of freedom, and twice the difference of the log likelihood values between the two models was compared with chi-square (χ^2) distribution. The comparison of M0 with M3 was used to evaluate the heterogeneity of ω among sites, while the other two comparisons (M2a versus M1a and M8 versus M7) were used to detect presence of sites evolving under positive selection. Amino acids under positive selection were identified using the Bayes empirical Bayes method (BEB) (Yang et al., 2005).

To test for possible rate heterogeneity of ω ratios across different branches, branch models of free-ratio model (Mf) and one-ratio model (M0) were compared using LRT method. Mf model allows each branch to have an independent ω , while the M0 allows all branches and sites to have a single ω ratio. Branch-site model was further performed to detect the role of positive selection acting on a specific branch (Zhang et al., 2005). For this analysis, the alternative model of branch-site model A (unfixed ω) and the corresponding null model (ω fixed to 1) were also compared using LRT. The branches of interest, such as within a lineage or a particular species, were used as foreground branches, and the remaining lineages or branches were used as background branches. Sites under positive selection in the foreground branch were identified using BEB method if the LRT is significant.

In addition, other four most likelihood methods implemented in the Datamonkey web server (see footnote 6) (Delpont et al., 2010), including Fixed Effects Likelihood (FEL), Single Likelihood Ancestral Counting (SLAC), Fast Unconstrained Bayesian AppRoximation (FUBAR), and Mixed Effects Model of Evolution (MEME), were also employed to test the impact of positive selection on *GME* and confirm the results of PAML.

RESULTS

Identification of *GME* Genes in Viridiplantae

BLAST results showed that *GME* genes are widely present in various plant species. A total of 111 *GME* homologous from 59 Viridiplantae species were obtained in the final data (Supplementary Table 1 and Supplementary Data 1, 2). These plant species covered a broad range of the major lineages of Viridiplantae, including dicotyledons, monocotyledons, gymnosperms, pteridophytes, bryophytes, and unicellular green algae. *GME* sequences of *A. deliciosa*, *A. rufa*, *A. erantina*, and *Picea sitchensis* were downloaded from NCBI database, while the rest of *GME* sequences were retrieved from Phytozome v12.1 database. The detailed information of the obtained plant *GME* sequences was provided in Supplementary Table 1.

All obtained *GME* candidates were checked to confirm the presence of conserved domains in their proteins sequences. The results of the NCBI conserved domain database showed that all the predicted *GME* proteins used in this study containing the NDA-binding domain. In addition to this, the conserved epimerase domain was also detected in all the predicted *GME* proteins by both the NCBI conserved domain search and the HMMER search. These results indicated that the obtained *GME* proteins have similar epimerase function.

Most Viridiplantae species have different *GME* gene copy numbers (Supplementary Table 1). In some species, such as *A. thaliana*, *Boechera stricta*, *Spirodela polyrhiza*, *M. polymorpha*, and *Volvox carteri*, only one *GME* copy was identified. However, in some other plant species, such as *Brassica rapa*, *Zea mays*, *P. sitchensis*, *Selaginella moellendorffii*, and *Physcomitrella patens*, more than two copies of *GME* were identified. Most of the plants used in this study contained only one or two *GME* copies, while the eudicot plant *Glycine max*, the monocot plant *Panicum virgatum*, the gymnosperm plant *P. sitchensis*, the lycophyte plant *S. moellendorffii*, and the bryophyte plant *P. patens* possessed 5, 5, 3, 9, 3 *GME* gene copies, respectively.

Recombination Test and Phylogenetic Analyses

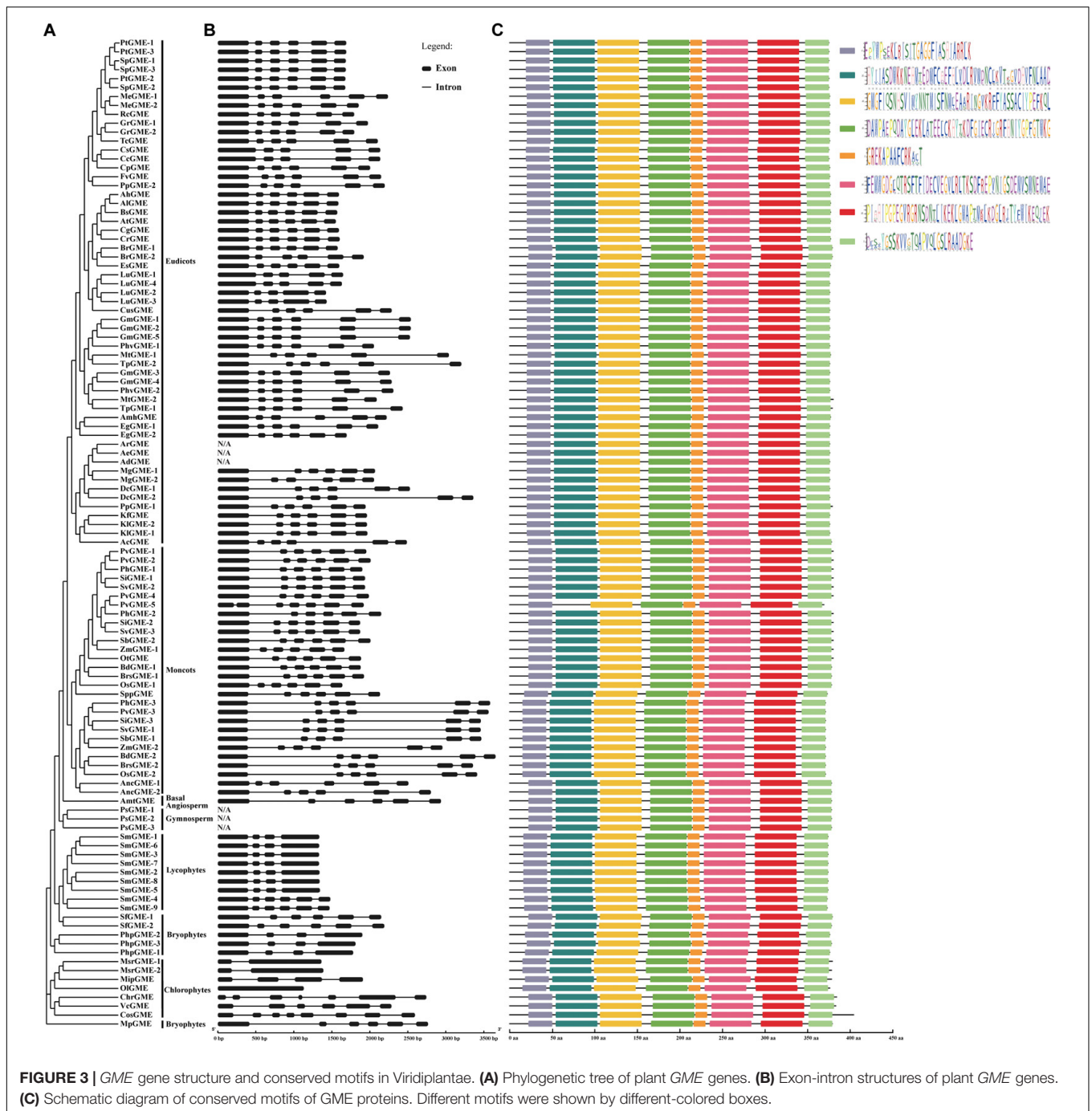
Gaps in more than 50% of the sequences in the nucleotide sequence alignment were removed. No evidence of recombination was found by GARD method implemented in Datamonkey website, suggesting the sequence alignment of the collected data could be used in subsequent phylogenetic reconstruction and molecular evolutionary analyses. The best-fit model of GTR+I+G was selected by AIC standard in MrModeltest and used in phylogenetic reconstruction analysis. Results showed that the Bayesian phylogenetic tree of plant *GME* was basically consistent with the evolutionary relationship of Viridiplantae (Figure 2). All of the angiosperm plant *GME* genes (including eudicots, monocots, and basal angiosperm) formed a single clade with high bootstrap support as showed in *GME* gene tree (Figure 2). Plant *GME* genes in gymnosperm, lycophytes, and chlorophytes also formed a single clade, respectively. However, *GME* genes of bryophytes did not form a monophyletic clade, and were positioned in three different clades (Figure 2).

Genomic Structure of *GME* Genes and Sequence Feature of *GME* Proteins

The exon-intron organization was examined to understand the structural diversity and the evolutionary relationship of plant *GME* genes. As the *GME* genomic DNA sequences of *A. deliciosa*, *A. rufa*, *A. erantina*, and *P. sitchensis* were unavailable right now, their exon-intron structures were not displayed. The exon-intron structures of plant *GME* genes were relatively conserved and most of the neighboring sequences in the phylogenetic tree tend to have similar gene structures according to the results (Figure 3). Except a few exceptions, the *GME* gene structures in the subgroup of angiosperms (including eudicots, monocots, and basal angiosperm) were highly conserved and almost all contained six exons in their coding regions. The *CsGME* (*Citrus*



FIGURE 2 | Phylogenetic relationships of *GME* gene in Viridiplantae. The phylogenetic tree of plant *GMEs* was constructed by the Bayesian analyses method based on the best fitted nucleotide substitution model of GTR+I+G. Posterior probabilities above 0.6 were labeled near the nodes. Accession number of each *GME* sequence was listed behind the species name. Branches identified under episodic diversifying selection by branch-site model were labeled using asterisk marks (★).



sinensis), *CcGME* (*Citrus clementina*), and *LuGME* (*Linum usitatissimum*) in eudicots subgroup contained five exons, while *PvGME-5* (*P. virgatum*) in monocots contained seven exons. It could be concluded that one intron of *CsGME*, *CcGME*, and *LuGME* may be lost during the evolution of these sequences by comparing their structure with others. For *PvGME-5*, the first exon appears truncated and joined with another; this sequence (369 aa) with seven exons is the shortest in this study (Figure 3). Two main types of exon-intron structures (four and six exons) were identified in the subgroups of lycophytes and bryophytes.

In the chlorophytes subgroup, the *GME* gene structures were not conserved, and various types of gene structures can be observed including one, two, four, six, seven, and eight exons. Except for *OIGME* (*Ostreococcus lucimarinus*) where no intron structure was observed, most Viridiplantae species contained at least one intron, and the intron number or length was variable among most species.

The full-length of most *GME* amino acid sequences was around 377 aa (Supplementary Table 1) and shared 62.6–100% pairwise protein sequence identity. The length of *GME* varied

from 369 to 447 aa (**Supplementary Table 1**), suggesting certain degree of insertion or deletion events may have occurred during the evolution of plant GME. To better understand the protein evolution, the conserved motifs of GME were investigated using MEME v4.12.0. Unlike the variation of exon-intron structures, the basic motifs of GME were strongly conserved throughout Viridiplantae (**Figure 3**). Except PvGME-5, which lost one motif and contained only seven motifs, all remaining protein sequences contained eight conserved motifs.

Molecular Evolution Analyses

Maximum-likelihood methods implemented in PAML, including specific site model, branch model, and branch-site model, were used to detect the selective pressures acting on plant GME genes. Three pairs of site-specific models were conducted to test positively selected sites on the evolution of plant GME. The maximum likelihood estimates of the ω values under the one-ratio model M0 for GME genes was 0.0287 (**Supplementary Table 2**), which approximates zero, indicating that the evolution of Viridiplantae GME genes was under strong purifying selection. The LRT test for model M0 versus M3 was statistically significant ($-2\Delta\ln L = 1861.68$, $p < 0.001$) (**Supplementary Table 2**), indicating that ω ratio was variable among amino acid sites along GME sequences. Then to determine the selective constraint among sites and identify potential sites under positive selection, the designated models were compared against their corresponding null models (i.e., M2a versus M1a and M8 versus M7). Both LRTs showed that the alternative models of M2a and M8 were not significantly better than null models of M1a and M7, respectively, and no amino acid site was detected to be under positive selection (**Supplementary Table 2**).

As positive selection may only influence some specific branches, branch-specific models of free-ratio model (Mf) and one-ratio model (M0) were compared to detect selective constraints among different branches. The LRT test showed that the Mf was significantly better than the M0 ($-2\Delta\ln L = 5197.214$, $p < 0.001$), suggesting that selective constraint level was heterogeneous among branches (**Supplementary Table 3**). The modified branch-site models were further used to detect selective pressure on different clades or branches and also identify sites influenced by positive selection. Firstly, the main lineages of eudicots, monocots, gymnosperm, lycophytes, and chlorophytes were used as foreground branch, respectively, while rest branches in the phylogenetic tree were used as background branches. According to the LRTs results, no significant differences were identified between the foreground branches and the background branches (**Supplementary Table 4**). Secondly, to test whether a particular branch was under positive selection during evolution process, each branch in the phylogenetic tree was labeled as foreground branch and the rest of the branches were treated as background branch. The LRTs identified 10 branches, including *A. thaliana*, *B. rapa* (BrGME-1), *B. rapa* (BrGME-2), *Chlamydomonas reinhardtii*, *Coccomyxa subellipsoidea*, *Medicago truncatula* (MtGME-2), *Micromonas pusilla*, *O. lucimarinus*, *V. carteri*, *M. polymorpha*, were significantly better fit to the data than their corresponding null models (**Supplementary Table 4**). And amino acid sites under

episodic positive selection were identified in each foreground branch with BEB posterior probabilities criterion of 95% (**Supplementary Table 4**). It was worth noting that five branches of *C. reinhardtii*, *C. subellipsoidea*, *M. pusilla*, *O. lucimarinus*, *V. carteri* belonged to chlorophytes subgroup, one branch of *M. polymorpha* belonged to bryophytes subgroup and four branches of *A. thaliana*, *B. rapa* (BrGME-1), *B. rapa* (BrGME-2), *M. truncatula* (MtGME-2) belonged to eudicots subgroup (**Figure 2**).

In addition, the FEL, SLAC, FUBAR, and MEME methods implemented in Datamonkey web server were also used to detect evidence of positive selection and verify the results of PAML. Results showed that no specific codons under positive selection were identified by methods of FEL, SLAC, and FUBAR. However, MEME identified 11 sites with evidence of episodic diversifying selection with significance level of $p < 0.1$ (**Supplementary Table 5**).

DISCUSSION

The reaction catalyzed by GME involves two pathways of plant AsA biosynthesis and is regarded to be a key link in plant AsA biosynthesis. Previous works about GME mainly focused on model plants and some cash crops, and found two copies of GME in tomato (Zhang et al., 2011) and only one copy of GME in most of other plants such as *Arabidopsis* (Wolucka et al., 2001), rice (Watanabe et al., 2006), kiwifruit (Li et al., 2013), and tea (Li et al., 2017). In this study, 111 GME homologous were obtained from 59 representative Viridiplantae species, which covered all the main lineages of plants including angiosperms, gymnosperms, lycophytes, bryophytes, and green algae. Except for a few plants, most Viridiplantae contained one or two GME homologous in the collected data set. Some plants contained more than three copies of GME, such as *S. moellendorffii* which contained as much as nine GME homologous (**Supplementary Table 1**). However, no signal of ancient whole genome duplication or polyploidy event was identified in *Selaginella* genome (Banks et al., 2011), and the mechanism of the formation of multi-copy of GME in these plants is still unclear. Rice genome encodes three homologs of GMPase (GDP-mannose pyrophosphorylase) which catalyzes the synthesis of GDP-D-mannose in L-galactose pathway. However, only two of the homologs involved in rice AsA biosynthesis and played different roles in the process (Qin et al., 2016). Two GME homologs in tomato participated in the biosynthesis of AsA and cell wall. However, they were specifically expressed in different organs or tissues, and had diversified roles in cell wall biosynthesis and development (Mounet-Gilbert et al., 2016). For the plants containing multiple homologs of GME, whether these homologs all participate in AsA biosynthesis and have same functions are still uncertain and needed to be verified in future works.

Most of the plant GME genes have similar structure. GME exon-intron structure was much conserved in eudicots and monocots, and most of the sequences contained six exon structures (**Figure 3**). The length and the arrangement of the exons in higher plants were also highly conserved. However, in

the basal subgroups of the phylogenetic tree, such as lycophytes, bryophytes, and chlorophytes, the structural types of *GME* gene varied greatly (Figure 3). For example, six types of exon-intron structures (one, two, four, six, seven, and eight exons) were identified in the subgroup of chlorophytes. These diverse gene structures indicated that *GME* gene structure may have undergone severe differentiation at the early diverging lineages of Viridiplantae, and the structure such as exon number was gradually fixed as plant evolution. Protein is the ultimate executor of the function of the gene. *GME* protein motif analyses showed that the protein structure was much conserved even in the basal subgroups of the phylogenetic tree where the gene structure was relatively diversified (Figure 3). The conserved motif of *GME* protein throughout the phylogenetic tree reflects the conserved and important function of the protein.

Stressful environments, such as high/low temperature, drought and salinity, can seriously affect normal growth and development of plants by inducing the over-production of reactive oxygen species (ROS). To adapt to these abiotic stresses, plants have evolved antioxidant systems consisting of enzymatic and non-enzymatic antioxidants to neutralize ROS and protect cells from oxidative damage. AsA is an efficient and major non-enzymatic antioxidant molecule and can improve plant tolerance to abiotic stresses by over-expressing AsA biosynthesis related genes or exogenously applying AsA (Akram et al., 2017). Therefore, increasing AsA level in plants is an effective way to enhance plant resistance to stresses. Different plants contain different concentrations of AsA, and understanding the molecular evolution mechanisms that causes the concentration variation will be helpful to increase AsA levels in plants. *GME* was regarded as a key gene in plant AsA biosynthesis pathway. However, *GME* alone had small effect on AsA pool in some plants, and it usually needed to cooperate with GDP-L-galactose phosphorylase (*GGP*) to control AsA biosynthesis (Mellidou and Kanellis, 2017). In this study, strong purify selection was found to be the predominant force in the evolution of *GME*, which suggesting functional importance of the gene. Nevertheless, a small subset of branches and sites under positive selection were identified by two episodic diversifying selection methods of branch-site model and MEME. Most interesting, 5 out of 10 branches under positive selection were located in the subgroup of green algae, which indicated that diversifying selection mainly occurred in the early diverging lineages of Viridiplantae.

Plant biosynthetic pathways usually involve many genes, which are in the network of protein–protein interactions or the interaction between products and substrates. Previous studies showed that the evolutionary relationships of genes along the metabolic pathway were different, and the evolutionary rate of

genes has a certain relationship with the position of genes in the pathway (Ramsay et al., 2009; Clotault et al., 2012; Nougé et al., 2014). L-galactose pathway is the major route for AsA biosynthesis in plants and genes involved in this pathway had been well documented (Akram et al., 2017). This study was the first attempt to illustrate the evolutionary pattern of the key gene *GME* which involved in L-galactose pathway. Molecular evolution analysis of other genes along the L-galactose pathway in future studies will be further improve our understanding of the evolutionary patterns of the pathway and the molecular mechanisms of AsA concentration difference among different plants.

CONCLUSION

The evolutionary pattern of plant AsA biosynthesis pathway gene *GME* was analyzed in large-scale for the first time in this study. Molecular evolutionary analyses showed that *GME* was under strong purifying selection, and only a few branches and sites under episodic diversifying selection were identified, suggesting episodic diversifying selection at a few branches and sites may play a role in the evolution of *GME*. The branches under positive selection mainly presented in the lineage of chlorophytes, where the *GME* gene structures were also the most diverse, and these results suggested that *GME* had undergone severe differentiation and diversifying selection in the early stage of Viridiplantae evolution.

AUTHOR CONTRIBUTIONS

JT and XX conceived and designed the project. JT, HW, and ZL collected the data and performed the experiments. JT and CH analyzed the data. JT wrote the paper.

FUNDING

This study was supported by grant from the science and technology research project of the education department of Jiangxi Province (GJJ150382).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.01293/full#supplementary-material>

REFERENCES

- Agius, F., Gonzalez-Lamothe, R., Caballero, J. L., Munoz-Blanco, J., Botella, M. A., and Valpuesta, V. (2003). Engineering increased vitamin C levels in plants by overexpression of a D-galacturonic acid reductase. *Nat. Biotechnol.* 21, 177–181. doi: 10.1038/nbt777
- Akram, N. A., Shafiq, F., and Ashraf, M. (2017). Ascorbic acid-A potential oxidant scavenger and its role in plant development and abiotic stress tolerance. *Front. Plant Sci.* 8:613. doi: 10.3389/fpls.2017.00613
- Anisimova, M., Nielsen, R., and Yang, Z. (2003). Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. *Genetics* 164, 1229–1236.

- Banks, J. A., Nishiyama, T., Hasebe, M., Bowman, J. L., Gribskov, M., dePamphilis, C., et al. (2011). The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science* 332, 960–963. doi: 10.1126/science.1203810
- Barth, C., De Tullio, M., and Conklin, P. L. (2006). The role of ascorbic acid in the control of flowering time and the onset of senescence. *J. Exp. Bot.* 57, 1657–1665. doi: 10.1093/jxb/erj198
- Bulley, S., and Laing, W. (2016). The regulation of ascorbate biosynthesis. *Curr. Opin. Plant Biol.* 33, 15–22. doi: 10.1016/j.pbi.2016.04.010
- Bulley, S. M., Rassam, M., Hoser, D., Otto, W., Schunemann, N., Wright, M., et al. (2009). Gene expression studies in kiwifruit and gene over-expression in *Arabidopsis* indicates that GDP-L-galactose guanyltransferase is a major control point of vitamin C biosynthesis. *J. Exp. Bot.* 60, 765–778. doi: 10.1093/jxb/ern327
- Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552. doi: 10.1093/oxfordjournals.molbev.a026334
- Clotault, J., Peltier, D., Soufflet-Freslon, V., Briard, M., and Geoffriau, E. (2012). Differential selection on carotenoid biosynthesis genes as a function of gene position in the metabolic pathway: a study on the carrot and dicots. *PLoS One* 7:e38724. doi: 10.1371/journal.pone.0038724
- Delpont, W., Poon, A. F. Y., Frost, S. D. W., and Kosakovsky Pond, S. L. (2010). Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* 26, 2455–2457. doi: 10.1093/bioinformatics/btq429
- Doskey, C. M., Buranasudja, V., Wagner, B. A., Wilkes, J. G., Du, J., Cullen, J. J., et al. (2016). Tumor cells have decreased ability to metabolize H₂O₂: implications for pharmacological ascorbate in cancer therapy. *Redox. Biol.* 10, 274–284. doi: 10.1016/j.redox.2016.10.010
- Finn, R. D., Clements, J., Arndt, W., Miller, B. L., Wheeler, T. J., Schreiber, F., et al. (2015). HMMER web server: 2015 update. *Nucleic Acids Res.* 43, W30–W38. doi: 10.1093/nar/gkv397
- Gallie, D. R. (2012). The role of L-ascorbic acid recycling in responding to environmental stress and in promoting plant growth. *J. Exp. Bot.* 64, 433–443. doi: 10.1093/jxb/ers330
- Gest, N., Gautier, H., and Stevens, R. (2013). Ascorbate as seen through plant evolution: the rise of a successful molecule? *J. Exp. Bot.* 64, 33–53. doi: 10.1093/jxb/ers297
- Gilbert, L., Alhagdog, M., Nunes-Nesi, A., Quemener, B., Guillon, F., Bouchet, B., et al. (2009). GDP-D-mannose 3,5-epimerase (GME) plays a key role at the intersection of ascorbate and non-cellulosic cell-wall biosynthesis in tomato. *Plant J.* 60, 499–508. doi: 10.1111/j.1365-3113.2009.03972.x
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186. doi: 10.1093/nar/gkr944
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.
- Huang, H., Wang, Y., Zhang, Z., Jiang, Z., and Wang, S. (2004). Actinidia germplasm resources and kiwifruit industry in China. *HortScience* 39, 1165–1172.
- Imai, T., Ban, Y., Terakami, S., Yamamoto, T., and Moriguchi, T. (2009). L-ascorbate biosynthesis in peach: cloning of six L-galactose pathway-related genes and their expression during peach fruit development. *Physiol. Plant.* 136, 139–149. doi: 10.1111/j.1399-3054.2009.01213.x
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kosakovsky Pond, S. L., Posada, D., Gravenor, M. B., Woelck, C. H., and Frost, S. D. (2006). GARD: a genetic algorithm for recombination detection. *Bioinformatics* 22, 3096–3098. doi: 10.1093/bioinformatics/btl474
- Kotchoni, S. O., Larrimore, K. E., Mukherjee, M., Kempinski, C. F., and Barth, C. (2009). Alterations in the endogenous ascorbic acid content affect flowering time in *Arabidopsis*. *Plant Physiol.* 149, 803–815. doi: 10.1104/pp.108.132324
- Li, H., Huang, W., Wang, G.-L., Wang, W.-L., Cui, X., and Zhuang, J. (2017). Transcriptomic analysis of the biosynthesis, recycling, and distribution of ascorbic acid during leaf development in tea plant (*Camellia sinensis* (L.) O. Kuntze). *Sci. Rep.* 7:46212. doi: 10.1038/srep46212
- Li, J., Cui, M., Li, M., Wang, X., Liang, D., and Ma, F. (2013). Expression pattern and promoter analysis of the gene encoding GDP-d-mannose 3',5'-epimerase under abiotic stresses and applications of hormones by kiwifruit. *Sci. Hortic.* 150, 187–194. doi: 10.1016/j.scienta.2012.11.008
- Liu, Y., Yu, L., and Wang, R. (2010). Level of ascorbic acid in transgenic rice for l-galactono-1,4-lactone dehydrogenase overexpressing or suppressed is associated with plant growth and seed set. *Acta Physiol. Plant.* 33, 1353–1363. doi: 10.1007/s11738-010-0669-5
- Lorence, A., Chevone, B. I., Mendes, P., and Nessler, C. L. (2004). myo-inositol oxygenase offers a possible entry point into plant ascorbate biosynthesis. *Plant Physiol.* 134, 1200–1205. doi: 10.1104/pp.103.033936
- Ma, Y., Chapman, J., Levine, M., Polireddy, K., Drisko, J., and Chen, Q. (2014). High-dose parenteral ascorbate enhanced chemosensitivity of ovarian cancer and reduced toxicity of chemotherapy. *Sci. Transl. Med.* 6:222ra218. doi: 10.1126/scitranslmed.3007154
- Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., et al. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 43, D222–D226. doi: 10.1093/nar/gku1221
- Mellado, M., Contreras, R. A., Gonzalez, A., Dennett, G., and Moenne, A. (2012). Copper-induced synthesis of ascorbate, glutathione and phytochelatins in the marine alga *Ulva compressa* (Chlorophyta). *Plant Physiol. Biochem.* 51, 102–108. doi: 10.1016/j.plaphy.2011.10.007
- Mellidou, I., and Kanellis, A. K. (2017). Genetic control of ascorbic acid biosynthesis and recycling in horticultural crops. *Front. Chem.* 5:50. doi: 10.3389/fchem.2017.00050
- Mounet-Gilbert, L., Dumont, M., Ferrand, C., Bournonville, C., Monier, A., Jorly, J., et al. (2016). Two tomato GDP-D-mannose epimerase isoforms involved in ascorbate biosynthesis play specific roles in cell wall biosynthesis and development. *J. Exp. Bot.* 67, 4767–4777. doi: 10.1093/jxb/erw260
- Nougué, O., Corbi, J., Ball, S., Manicacci, D., and Tenailon, M. (2014). Molecular evolution accompanying functional divergence of duplicated genes along the plant starch biosynthesis pathway. *BMC Evol. Biol.* 14:103. doi: 10.1186/1471-2148-14-103
- Nylander, J. (2008). *MrModeltest v2.3. Program distributed by the author. Evolutionary Biology Centre, Uppsala University website.* Available at: <https://github.com/nylander/MrModeltest2>.
- Obinger, C., Regelsberger, G., Pircher, A., Strasser, G., and Peschek, G. A. (1998). Scavenging of superoxide and hydrogen peroxide in blue-green algae (cyanobacteria). *Physiol. Plant.* 104, 693–698. doi: 10.1034/j.1399-3054.1998.1040425.x
- Paciolla, C., and Tommasi, F. (2003). The ascorbate system in two bryophytes: *Brachythecium velutinum* and *Marchantia polymorpha*. *Biol. Plant.* 47, 387–393.
- Posada, D., and Crandall, K. A. (2002). The effect of recombination on the accuracy of phylogeny estimation. *J. Mol. Evol.* 54, 396–402. doi: 10.1007/s00239-001-0034-9
- Qi, T., Liu, Z., Fan, M., Chen, Y., Tian, H., Wu, D., et al. (2017). GDP-D-mannose epimerase regulates male gametophyte development, plant growth and leaf senescence in *Arabidopsis*. *Sci. Rep.* 7:10309. doi: 10.1038/s41598-017-10765-5
- Qin, H., Deng, Z., Zhang, C., Wang, Y., Wang, J., Liu, H., et al. (2016). Rice GDP-mannose pyrophosphorylase OsVTC1-1 and OsVTC1-3 play different roles in ascorbic acid synthesis. *Plant Mol. Biol.* 90, 317–327. doi: 10.1007/s11103-015-0420-0
- Ramsay, H., Rieseberg, L. H., and Ritland, K. (2009). The correlation of evolutionary rate with pathway position in plant terpenoid biosynthesis. *Mol. Biol. Evol.* 26, 1045–1053. doi: 10.1093/molbev/msp021
- Ronquist, F., and Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574. doi: 10.1093/bioinformatics/btg180
- Shiu, C. T., and Lee, T. M. (2005). Ultraviolet-B-induced oxidative stress and responses of the ascorbate-glutathione cycle in a marine macroalga *Ulva fasciata*. *J. Exp. Bot.* 56, 2851–2865. doi: 10.1093/jxb/eri277
- Shriner, D., Nickle, D. C., Jensen, M. A., and Mullins, J. I. (2003). Potential impact of recombination on sitewise approaches for detecting positive natural selection. *Genet. Res.* 81, 115–121. doi: 10.1017/s0016672303006128
- Siow, R. S., Teoh, S., Teo, S. S., Shukor, M. Y. B. A., Phang, S. M., and Ho, C. L. (2013). Molecular cloning and characterization of GDP-mannose-3',5'-epimerase from *Gracilaria changii*. *J. Appl. Phycol.* 25, 1309–1318. doi: 10.1007/s10811-013-9987-5
- Sun, S. Q., He, M., Cao, T., Yusuyin, Y., Han, W., and Li, J. L. (2010). Antioxidative responses related to H₂O₂ depletion in *Hypnum plumaeforme* under the

- combined stress induced by Pb and Ni. *Environ. Monit. Assess.* 163, 303–312. doi: 10.1007/s10661-009-0835-7
- Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34, W609–W612. doi: 10.1093/nar/gkl315
- Talla, S., Riazunnisa, K., Padmavathi, L., Sunil, B., Rajsheel, P., and Raghavendra, A. S. (2011). Ascorbic acid is a key participant during the interactions between chloroplasts and mitochondria to optimize photosynthesis and protect against photoinhibition. *J. Biosci.* 36, 163–173. doi: 10.1007/s12038-011-9000-x
- Tel-or, E., Huflejt, M. E., and Packer, L. (1986). Hydroperoxide metabolism in cyanobacteria. *Arch. Biochem. Biophys.* 246, 396–402.
- Voxeur, A., Gilbert, L., Rihouey, C., Driouich, A., Rothan, C., Baldet, P., et al. (2011). Silencing of the GDP-D-mannose 3,5-epimerase affects the structure and cross-linking of the pectic polysaccharide rhamnogalacturonan II and plant growth in tomato. *J. Biol. Chem.* 286, 8014–8020. doi: 10.1074/jbc.M110.198614
- Wang, P., Yin, L., Liang, D., Li, C., Ma, F., and Yue, Z. (2011). Delayed senescence of apple leaves by exogenous melatonin treatment: toward regulating the ascorbate–glutathione cycle. *J. Pineal Res.* 53, 11–20. doi: 10.1111/j.1600-079X.2011.00966.x
- Watanabe, K., Suzuki, K., and Kitamura, S. (2006). Characterization of a GDP-D-mannose 3',5'-epimerase from rice. *Phytochemistry* 67, 338–346. doi: 10.1016/j.phytochem.2005.12.003
- Wheeler, G. L., Jones, M. A., and Smirnov, N. (1998). The biosynthetic pathway of vitamin C in higher plants. *Nature* 393, 365–369. doi: 10.1038/30728
- Wolucka, B. A., and Montagu, M. V. (2007). The VTC2 cycle and the de novo biosynthesis pathways for vitamin C in plants: an opinion. *Phytochemistry* 68, 2602–2613.
- Wolucka, B. A., Persiau, G., Van Doorselaere, J., Davey, M. W., Demol, H., Vandekerckhove, J., et al. (2001). Partial purification and identification of GDP-mannose 3',5'-epimerase of *Arabidopsis thaliana*, a key enzyme of the plant vitamin C pathway. *Proc. Natl. Acad. Sci. U.S.A.* 98, 14843–14848. doi: 10.1073/pnas.011578198
- Wolucka, B. A., and Van Montagu, M. (2003). GDP-mannose 3',5'-epimerase forms GDP-L-gulose, a putative intermediate for the de novo biosynthesis of vitamin C in plants. *J. Biol. Chem.* 278, 47483–47490. doi: 10.1074/jbc.M309135200
- Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088
- Yang, Z., Nielsen, R., Goldman, N., and Pedersen, A.-M. K. (2000). Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155, 431–449.
- Yang, Z., Wong, W. S., and Nielsen, R. (2005). Bayes empirical bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* 22, 1107–1118. doi: 10.1093/molbev/msi097
- Ye, N., Zhu, G., Liu, Y., Zhang, A., Li, Y., Liu, R., et al. (2012). Ascorbic acid and reactive oxygen species are involved in the inhibition of seed germination by abscisic acid in rice seeds. *J. Exp. Bot.* 63, 1809–1822. doi: 10.1093/jxb/err336
- Yun, J., Mullarky, E., Lu, C., Bosch, K. N., Kavalier, A., Rivera, K., et al. (2015). Vitamin C selectively kills KRAS and BRAF mutant colorectal cancer cells by targeting GAPDH. *Science* 350, 1391–1396. doi: 10.1126/science.aaa5004
- Zhang, C., Liu, J., Zhang, Y., Cai, X., Gong, P., Zhang, J., et al. (2011). Overexpression of SIGMEs leads to ascorbate accumulation with enhanced oxidative stress, cold, and salt tolerance in tomato. *Plant Cell Rep.* 30, 389–398. doi: 10.1007/s00299-010-0939-0
- Zhang, C., Ouyang, B., Yang, C., Zhang, X., Liu, H., Zhang, Y., et al. (2013). Reducing AsA leads to leaf lesion and defence response in knock-down of the AsA biosynthetic enzyme GDP-D-Mannose pyrophosphorylase gene in tomato plant. *PLoS One* 8:e61987. doi: 10.1371/journal.pone.0061987.g001
- Zhang, J., Nielsen, R., and Yang, Z. (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol. Biol. Evol.* 22, 2472–2479. doi: 10.1093/molbev/msi237

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Tao, Wu, Li, Huang and Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.