

RESEARCH ARTICLE

Open Access



# Genome-wide identification, characterization and expression analysis of MATE family genes in apple (*Malus × domestica* Borkh)

Weihan Zhang<sup>1†</sup>, Liao Liao<sup>2†</sup>, Jinsheng Xu<sup>1</sup>, Yuepeng Han<sup>2</sup> and Li Li<sup>1,3\*</sup> 

## Abstract

**Background:** As an important group of the multidrug efflux transporter family, the multidrug and toxic compound extrusion (MATE) family has a wide range of functions and is distributed in all kingdoms of living organisms. However, only two MATE genes in apple have been analyzed and genome-wide comprehensive analysis of MATE family is needed.

**Results:** In this study, a total of 66 MATE (*MdMATE*) candidates encoding putative MATE transporters were identified in the apple genome. These *MdMATE* genes were classified into four groups by phylogenetic analysis with MATE genes in *Arabidopsis*. Synteny analysis reveals that whole genome duplication (WGD) and segmental duplication events played a major role in the expansion of MATE gene family in apple. *MdMATE* genes show diverse expression patterns in different tissues/organs and developmental stages. Analysis of *cis*-regulatory elements in *MdMATE* promoter regions indicates that the function of *MdMATE* genes is mainly related to stress response. Besides, the changes of gene expression levels upon different pathogen infections reveal that *MdMATE* genes are involved in biotic stress response.

**Conclusions:** In this work, we systematically identified *MdMATE* genes in apple genome using a set of bioinformatics approaches. Our comprehensive analysis provided valuable resources for improving disease resistance in apple and further functional characterization of MATE genes in other species.

**Keywords:** Apple, MATE, Gene family, Expression profile, Disease resistance

## Background

The multidrug and toxic compound extrusion (MATE) protein belongs to a multidrug efflux transporter family, which plays a role in transporting multiple kinds of substrates, such as secondary metabolites and

phytohormones [1]. The process of detoxification mainly involves four transporter families: ATP-binding cassette superfamily (ABC), resistance/nodulation/division family (RND), small gene multidrug resistance family (SMR), and major facilitator superfamily (MFS) [2]. Members of the ABC superfamily are considered to be primary transporters, and ATP provides energy during transport [3]. The transmembrane transport of primary transporters is mainly carried out by releasing energy [4]. The members of RND, SMR and MFS are secondary transporters and mainly use the electrochemical penetration potential

\* Correspondence: [li.li@mail.hzau.edu.cn](mailto:li.li@mail.hzau.edu.cn)

<sup>†</sup>Weihan Zhang and Liao Liao contributed equally to this work.

<sup>1</sup>Hubei Key Laboratory of Agricultural Bioinformatics, College of Informatics, Huazhong Agricultural University, Wuhan 430070, People's Republic of China

<sup>3</sup>Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan 430070, People's Republic of China

Full list of author information is available at the end of the article



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

caused by the difference between the internal and external material concentration of the membrane to implement transport process [4]. In these secondary transporter families, numerous MFS family members with different functions are widely distributed in both higher and lower organisms. Most MFS superfamily members contained 12 membrane alpha-helix and use cations ( $H^+$  or  $Na^+$  ions) electrochemical gradients to drive substrate export [5, 6]. In 1998, a new multidrug efflux system named NorM that can increase the efflux of norfloxacin in the absence of a multidrug efflux system in the host was found in *E. coli*. Meanwhile, a high sequence homology multidrug efflux protein was also found and named YdhE [7]. Although the structure of NorM is similar to the MFS superfamily that has 12 transmembrane regions (TMs), both NorM and YdhE proteins have no homology with members of the MFS superfamily, nor a specific signal sequence for this family [8]. Therefore, the new secondary transporter family, NorM and YdhE, become the fifth kind of transporter family which is named as multidrug and toxic compound extrusion (MATE) [9].

Previous study has shown that MATE transporters are widely distributed in all kingdoms of living organisms [1]. The protein length of most MATE transporters is 400–700 amino acids, containing 8–12 transmembrane domains [10, 11]. The first plant MATE transporter was identified in *Arabidopsis* named *A. thaliana aberrant lateral root formation 5* (*AtALF5*), which is associated with roots epidermal cells development and toxic compounds export [12]. In recent years, several studies have been conducted to characterize the function of MATE proteins in the model organism *Arabidopsis* and have shown that MATE proteins have various functions. For example, the *TRANSPARENT TESTA 12* (*TT12*) encodes a MATE transporter which can mediate anthocyanin transport [13] and the *AtDTX1* (*DETOXIFICATION 1*) is the first multi-specific MATE transporter control the export of toxic compounds from the cytoplasm [10]. In contrast to the mammal, plant genomes carry a larger number of MATE genes [14]. For instance, 56 in *Arabidopsis* [10], 48 in potato [6], 49 in maize [15], 45 in rice [16] and 117 in soybean [17]. Extensive research on MATE gene family suggest diverse functions, which are involved in stress responses, secondary metabolite and phytohormone transport, plant growth and development [18, 19]. However, the function and status of MATE gene family in fruit trees have rarely been reported.

Apple (*Malus × domestica* Borkh.) is one of the most economically important and popular perennial fruit crops in temperate regions. It is highly susceptible to infection by pathogen and greatly affect quality and yield. The development of resistance in pathogens and the large investment of farmers in fungicides have become

an important issue [20]. A previous study showed that the expression levels of MATE gene family are involved in pathogen susceptibility [21]. However, as far as we know only two *MATE* genes (*MdMATE1* and *MdMATE2*) in apple were analyzed, which are homologs of *TT12* in *Arabidopsis* and participate in proanthocyanidins accumulating in cells and flavonoid transport [22]. A comprehensive analysis of the MATE gene family is needed in the apple genome.

In this study, we identified 66 *MATE* genes (*MdMATE*) in apple genome and conducted comprehensive analysis regarding their phylogenetic relationship, synteny, gene structure, evolution, expression in different tissues/organs and developmental stages, and *cis*-elements. Synteny analysis shows that WGD/segmental duplication events played a major role in the expansion of MATE gene family in apple. Expression analysis suggests that *MdMATE* genes have specific expression patterns in different tissues/organs and developmental stages. Correlation analysis reveals that the functions of duplicated gene pairs may have divergence. We also found that *MdMATE* genes are involved in biotic stress response. Our study provides clues for further functional studies of MATE genes in plants and improvement of disease resistance in apple.

## Results

### Genome-wide identification of MATE transporters in apple

A total of 66 genes encoding MATE transporters were identified in the apple reference genome GDDH13 version 1.1 [23] after homologous sequence alignment with *AtMATE* genes from *A. thaliana* and manual filtering (see Methods) (Additional file 1: Table S1). Since two MATE transporters, *MdMATE1* and *MdMATE2*, have been reported in previous study [22], we renamed the other 64 transporters as *MdMATE3* – *MdMATE66* based on their physical locations (Additional file 1: Table S2).

The basic properties including the length of protein sequence, theoretical isoelectric point (pI), molecular weight (MW) and subcellular localization were analyzed to further characterize the *MdMATE* proteins (Additional file 1: Table S2). The 66 *MdMATE* proteins consist of 406 to 712 amino acids in length, which is quite similar to MATE proteins in *Arabidopsis* (400–700 amino acids) [10], but different from soybean (80–593 amino acids) [17], *Populus* (120–608 amino acids) [24] and rice (469–575 amino acids) [16]. The MW and pI ranges from 45.20 to 77.02 kDA and 5.05 to 9.54, respectively. Subcellular location prediction results include plasma membrane (56), chloroplast (5), cytoplasm (2), vacuole (2), and endoplasmic reticulum (1).

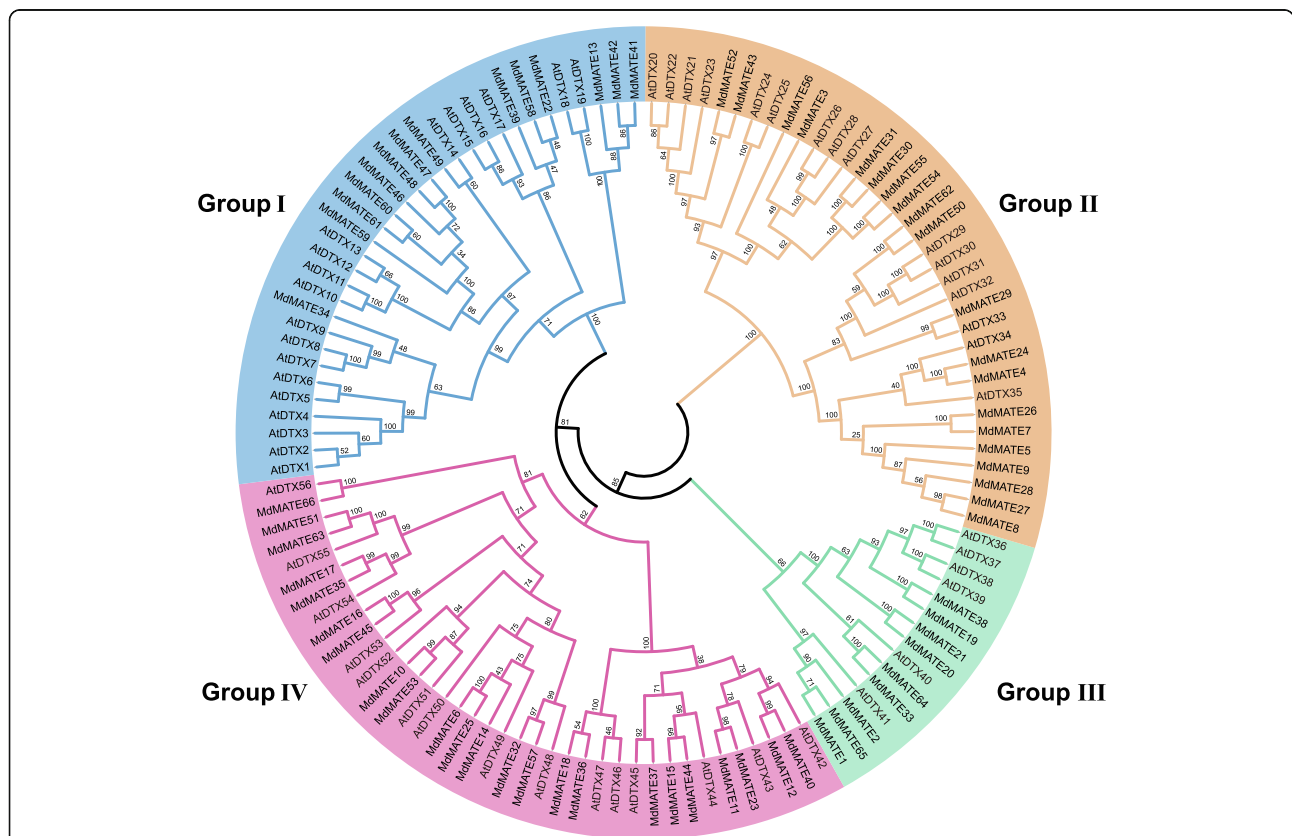
**Phylogenetic, gene structures and motif composition analysis of the MATE family in apple**

Using the full-length protein sequences of the 122 MATE transporters, including 56 from *Arabidopsis* and 66 identified in apple, we constructed a maximum likelihood (ML) phylogenetic tree (Fig. 1). The 66 MdMATE transporters were divided into four groups (I, II, III and IV), which is in accordance with what was reported for MATE transports in *Arabidopsis* [10]. A neighbor-joining (NJ) tree was also constructed to validate the phylogenetic relationship and show a similar grouping mode as ML tree (Additional file 2: Fig. S1). The group sizes vary from 9 to 23. Group I, II, III and IV contain 14, 20, 9 and 23 MdMATE transporters, respectively. We further aligned MdMATE transporters to construct an individual phylogenetic tree by the same method and parameters and got consistent clustering patterns (Fig. 2A).

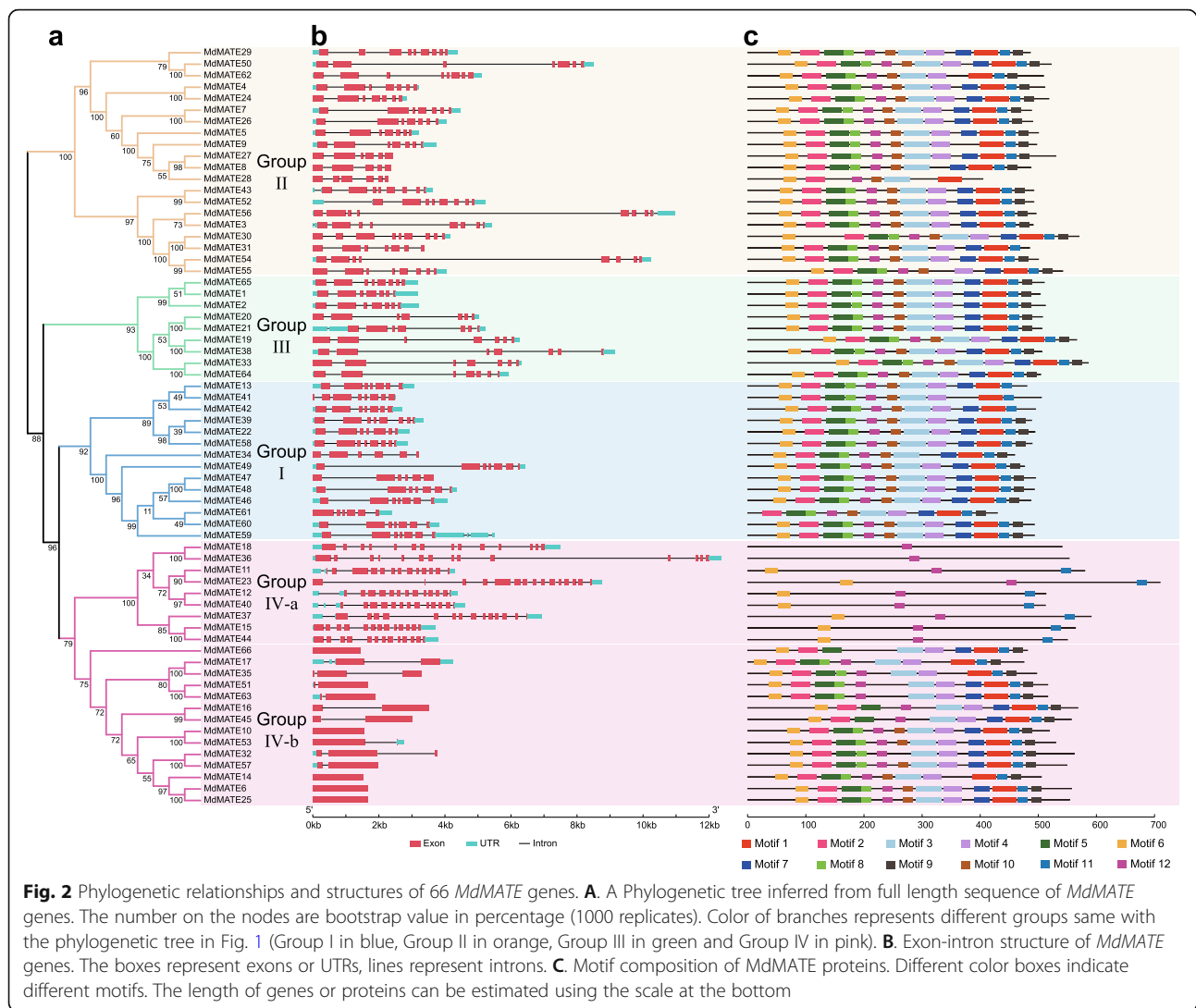
To gain more insight into the diversity of the MATE family in apple, we examined the exon-intron organization of all the identified *MdMATE* genes (Fig. 2B). The *MdMATE* genes display one to fifteen exons (5 with one exon, 7 with two exons, 2 with three exons,

one with five exons, 5 with six exons, 24 with seven exons, 11 with eight exons, and 11 with nine or more exons). The length of exons ranging from 9 (*MdMATE3*) to 1680 (*MdMATE6*) bp. The fewest number of exons are observed in *MdMATE6*, *MdMATE25*, *MdMATE14*, *MdMATE10*, and *MdMATE66*. These five genes all have only one exon and clustered in group IV. Additionally, 9 genes have more than thirteen exons. Overall, MATE gene family in apple showed complex gene structures with varying exons and lengths.

Motifs of the 66 MdMATE proteins were identified by MEME [25] to analyze the function and/or structural roles of highly conserved amino acid residues in active proteins [26]. A total of 12 conserved motifs, designated as motif 1 to motif 12, were identified (Fig. 2C and Additional file 2: Fig. S2). Most MdMATE proteins have similar motifs within the same group. Of these, motif 12 was found in all the MdMATE proteins except for *MdMATE66*. Interestingly, in group IV, some MdMATE proteins have fewer motifs ( $\leq 3$ ), but more exons ( $\geq 13$ ) than other genes. Based on the structure and motif differences in group IV, the group IV we further divided into two subgroups: Group IV-a and Group IV-b (Fig.



**Fig. 1** The unrooted maximum likelihood phylogenetic tree of MATE family members in apple and *Arabidopsis*. The different colors indicate different groups (Group I in blue, Group II in orange, Group III in green and Group IV in pink). ‘MdMATE’ represents MATE members from apple, ‘AtDXT’ represents MATE members from *Arabidopsis*. Numbers on the nodes are bootstrap values in percentage (1000 replicates)



2A). The group IV-a with more exons but fewer motifs, while the group IV-b is in contrast. These observations are consistent with studies of MATE gene family in other species such as soybean [17], cotton [27] and flax [28].

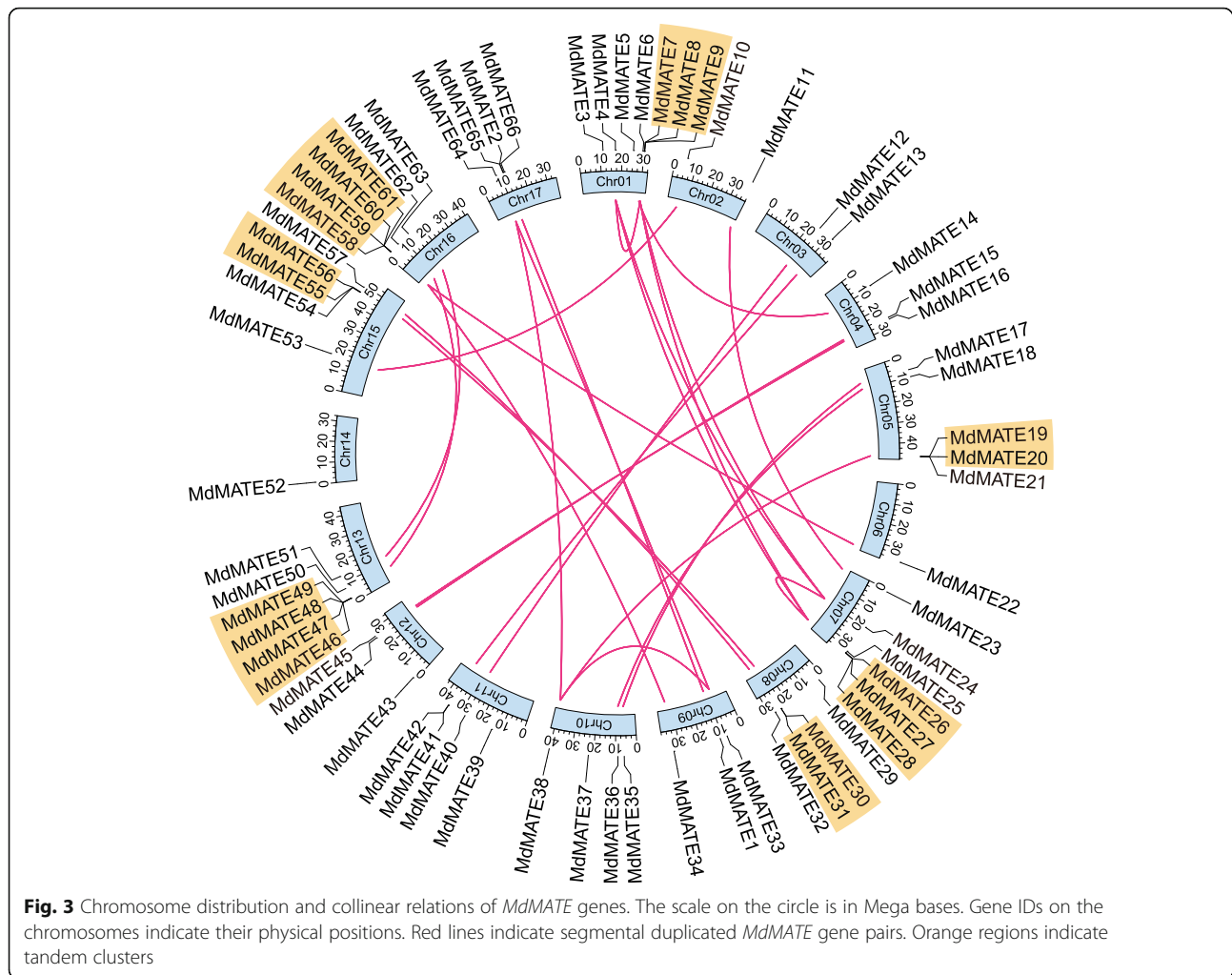
### Chromosomal location and evolution of apple MATE transporters

The distribution of *MdMATE* genes is uneven on the 17 chromosomes of apple genome ( $2n = 34$ ) (Fig. 3). Chromosome 1 contains the highest number of *MdMATE* genes (7), whereas both chromosomes 6 and 14 contain only one gene. There is no significant correlation between chromosome length and *MdMATE* gene number (Spearman's  $\rho = 0.317$ ,  $P = 0.2157$ ). Additionally, majority of these *MdMATE* genes are located on the chromosome arms, which are the regions with relative high recombination rate [29].

Gene duplications are considered to be one of the major driving forces in the evolution of genomes and expansions of the gene families [30, 31]. Whole genome duplication, segmental duplication and tandem duplication are the major causes of gene family expansion in plants [32]. We detected the duplicated events for *MdMATE* genes by MCscanX [33]. As shown in Figure 3, 26 gene pairs with 39 (59%) genes were identified as WGD/segmental duplication, while 20 (30%) tandem duplicated genes were identified within 7 tandem duplicated gene clusters (Additional file 1: Table S3). These results indicate that MATE family in apple expands mainly by gene duplications, with WGD/segmental duplication as the driving force.

### Syntenic analysis of MATE family in green plants

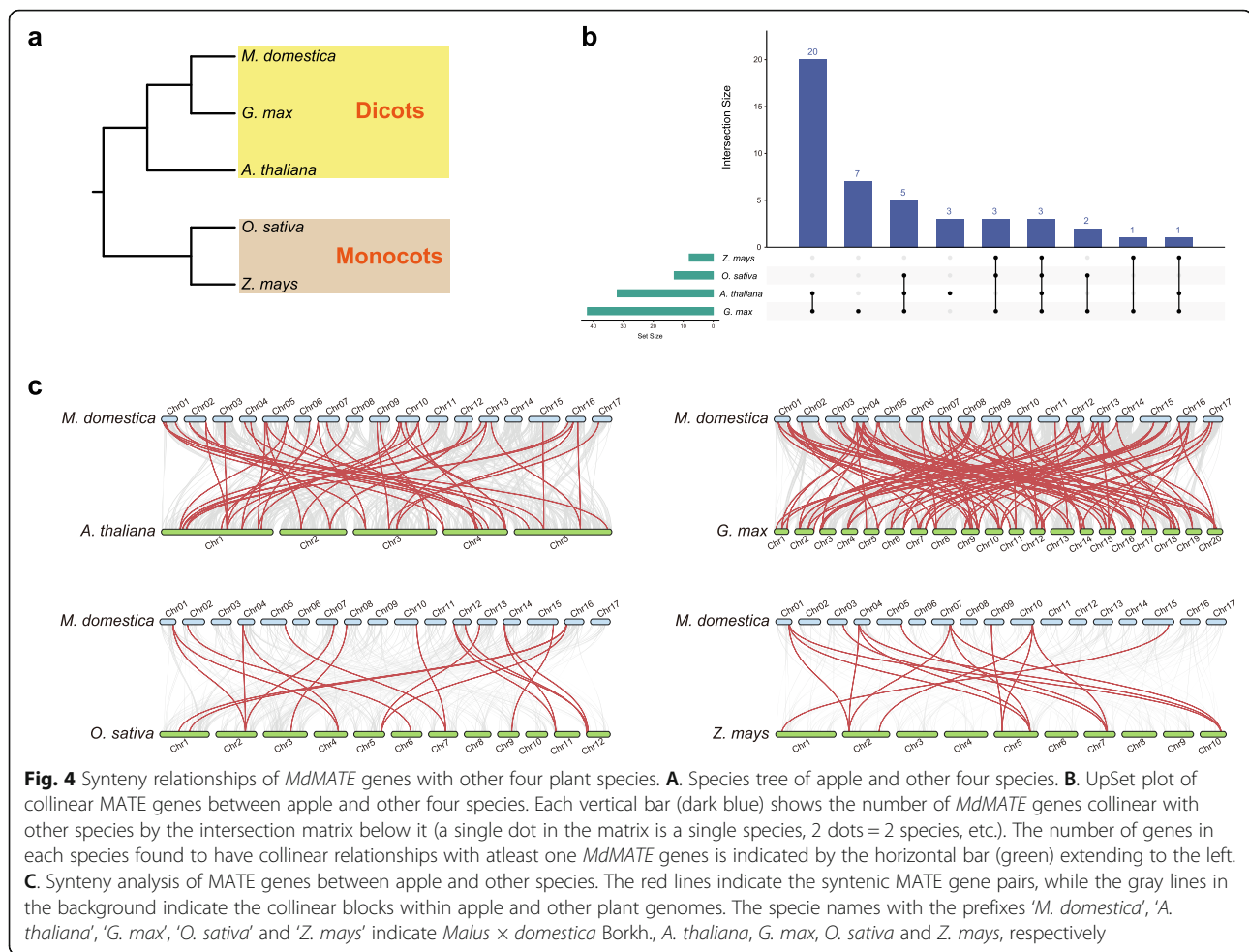
As an integral membrane protein involved in a diverse array of functions, MATE family of transporters are abundant in plants [11]. We constructed four



comparative syntenic maps of apple associated with four representative green plants to further investigate the phylogenetic mechanisms of MATE gene family in apple, including two dicots (*Arabidopsis* and soybean) and two monocots (maize and rice) (Fig. 4A). Among *MdMATE* genes, 68.2% (45 of 66 genes) were found associated with at least one collinear gene pair (Fig. 4B). A total of 42 *MdMATE* genes showed collinear relationship with 56 *GmMATE* genes in soybean, while 32 *MdMATE* genes showed collinear relationship with 26 *AtDTX* genes in *Arabidopsis*. In monocots, however, only 13 and 8 *MdMATE* genes collinear with 12 MATE genes in rice and 8 MATE genes in maize, respectively (Fig. 4C and Additional file 1: Table S4). Thus, the *MdMATE* genes show more collinear gene pairing with dicots than monocots.

For the collinear gene pairs between apple and soybean, which shows the most collinear relationships, 31 *MdMATE* genes are associated with two or more *GmMATE* genes. The *MdMATE14*, with 8 collinear *GmMATE* genes, is the gene with the most collinear

relationships. In contrast, for *Arabidopsis*, most *MdMATE* genes associated with one or two *AtDTX* genes. Only *MdMATE38* associated three *AtDTX* genes. This might result from two whole genome duplication events in soybean [34]. Interestingly, 30.3% (20 of 66 genes) of *MdMATE* genes have collinear gene pairs in both *Arabidopsis* and soybean, but no collinear gene pairs in monocots (Fig. 4B). For example, *MdMATE1* collinear with *AtDTX41* and *GmMATE81*, but shows no collinear gene in maize or rice. This is similar with observation of WRKY family in pineapple in the sense that these MATE orthologous gene pairs appear after the divergence of dicotyledonous and monocotyledonous plants [31]. Notably, three *MdMATE* genes, *MdMATE14*, *MdMATE19*, and *MdMATE38*, have collinear relationships between apple and all of the other four species, indicating that these orthologous pairs are conserved and may already exist before the ancestral divergence. These collinear gene pairs between apple and other species may be valuable for elucidating the evolution of MATE genes in green plants.



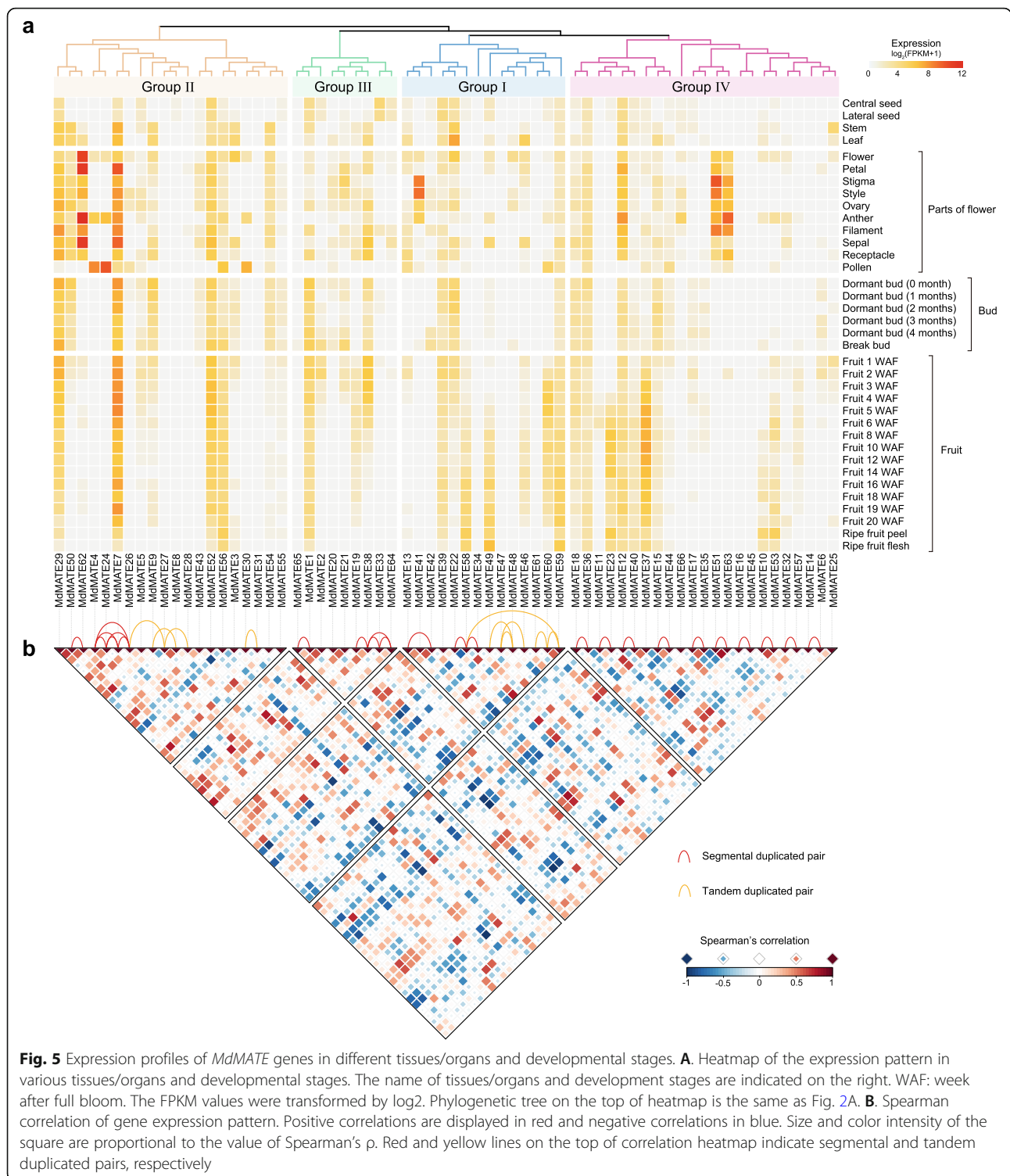
To better understand the different selective constraints on MATE gene family, the Ka/Ks ratios of the MATE gene pairs within apple and between apple and the other four plants were calculated (Additional file 1: Table S4). All segmental and tandem duplicated *MdMATE* gene pairs and the orthologous MATE gene pairs had Ka/Ks < 1, suggesting that the MATE gene family in apple have experienced strong purifying selective pressure during the evolution.

#### Expression patterns of *MdMATE* genes in different apple tissues/organs and developmental stages

To dissect the expression patterns of *MdMATE* genes in various tissues/organs and developmental stages, a total of 36 expression profiles of 66 *MdMATE* genes were obtained from Apple Multi-Dimensional Omics Database (AppleMDO) [35] (Fig. 5A and Additional file 1: Table S5). The *MdMATE* genes with Fragments Per Kilobase per Million (FPKM) values less than 1 in all 36 expression profiles are considered to be barely expressed [4]. Thus, 11 genes are not expressed in all 36 tissues/organs and developmental stage profiles. The remaining 55

genes are expressed in at least one profile. Among them, *MdMATE1* is expressed in all developmental stages of fruit flesh and peel, while *MdMATE2* shows expression level declining from earlier to later developmental stages. This is consistent to the real-time RT-PCR results published by previously study [22].

Among the 55 expressed genes, *MdMATE7* and *MdMATE36* showed constitutive expression (FPKM > 1 in all tissues/organs and developmental stages). It suggests that these two genes may involve apple growth. Some gene expressions are tissue-specific. For example, *MdMATE24* has high expression in parts of flower, especially in pollen. *MdMATE23* expressed mainly in fruit and peel. While *MdMATE62* is highly expressed in parts of flower except for pollen, *MdMATE33* is highly expressed in central and lateral seed. The expression level of some genes gradually increases as the fruit matures (eg. *MdMATE23*, *MdMATE49* and *MdMATE58*) and some genes were show the other way (eg. *MdMATE2*, *MdMATE5* and *MdMATE38*). It suggests their putative roles during apple fruit development. These results show that most *MdMATE* genes have



**Fig. 5** Expression profiles of *MdMATE* genes in different tissues/organs and developmental stages. **A.** Heatmap of the expression pattern in various tissues/organs and developmental stages. The name of tissues/organs and developmental stages are indicated on the right. WAF: week after full bloom. The FPKM values were transformed by  $\log_2$ . Phylogenetic tree on the top of heatmap is the same as Fig. 2A. **B.** Spearman correlation of gene expression pattern. Positive correlations are displayed in red and negative correlations in blue. Size and color intensity of the square are proportional to the value of Spearman's  $\rho$ . Red and yellow lines on the top of correlation heatmap indicate segmental and tandem duplicated pairs, respectively

dynamic expressions in different tissues/organs or developmental stages. It can help us to explore the functional diversity of *MdMATE* genes in apple.

After gene duplication, the divergence of expression pattern of two copies is considered to be an important aspect in their functional differentiation [36, 37]. In the

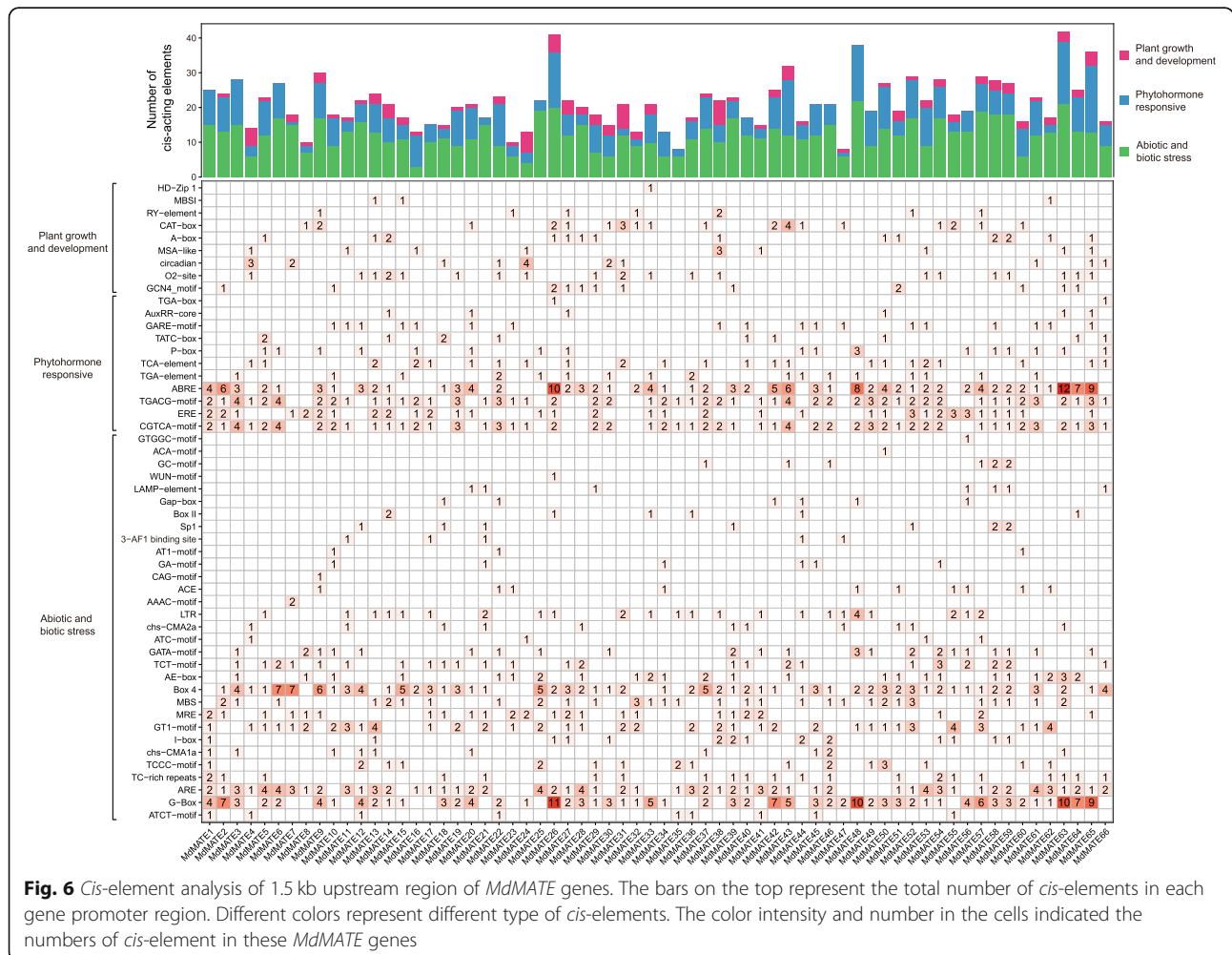
context of gene expression, it is straightforward to observe the function divergence after gene duplication [38]. Hence, correlation coefficient was calculated for the expression pattern of *MdMATE* gene pairs (Fig. 5B). The results show that genes in same group could differ considerably in their expression pattern, similar as before

[39]. Among the 39 duplicated pairs, *MdMATE4/MdMATE26*, *MdMATE10/MdMATE53*, *MdMATE17/MdMATE35*, *MdMATE50/MdMATE62*, *MdMATE51/MdMATE63*, *MdMATE58/MdMATE59*, and *MdMATE59/MdMATE60* showed significant positive correlations (Spearman's  $\rho > 0.60$ ,  $P < 0.05$ ), and only one gene pair *MdMATE22/MdMATE58* showed negative correlations (Spearman's  $\rho = -0.62$ ,  $P < 0.05$ ). As novel expression patterns may also only occur in specific organs, we further investigate the correlation of expression patterns in three organs (flower, bud and fruit) [40] (Additional file 2: Fig. S3). Intriguingly, some duplicated pairs showed organ-specific correlation (*MdMATE38/MdMATE64* in flower, *MdMATE13/MdMATE42* in bud and *MdMATE1/MdMATE65* in fruit) (Additional file 2: Fig. S3). Additionally, in fruit, genes in same group showed higher than overall correlation. Almost all members in group III (green) showed positive correlation. In summary, these results indicated that the functions of *MdMATE* genes in apple tissues/organs and developmental stages may be widely correlated

and varied. These paralogs with tissue-specific expression pattern has long been regarded as a precursor of future evolution which may contribute to phenotypic variation [36, 41].

**Cis-acting regulatory element analysis**

Gene transcription in plants is regulated by the *cis*-acting regulatory elements and transcription factors [42]. To identify putative *cis*-acting elements in the promoter region, we scanned the 1.5 kb upstream regions of transcriptional start site (TSS). All the candidate *MdMATE* gene promoters possessed typical TATA and CAAT boxes which are the core *cis*-acting element in eukaryotic promoter and enhancer regions (Additional file 1: Table S6). Other *cis*-acting regulatory elements were grouped into three main types based on their functional annotation: plant growth and development, phytohormone responsive and abiotic and biotic stress (Fig. 6). The abiotic and biotic stress group had the most number of regulatory elements, such as G-box, MBS and LTR, which were responsive to light inducible, drought



**Fig. 6** *Cis*-element analysis of 1.5 kb upstream region of *MdMATE* genes. The bars on the top represent the total number of *cis*-elements in each gene promoter region. Different colors represent different type of *cis*-elements. The color intensity and number in the cells indicated the numbers of *cis*-element in these *MdMATE* genes



inducible and low-temperature stress, respectively. Many abiotic and biotic stress elements were observed in the promoter region of *MdMATE* genes, revealing that *MdMATE* genes play important roles on the stress response. The followed group is phytohormone responsive which has 11 elements, the ERE, ABRE, TGACG-motif and CGTCA-motif widely distributed among *MdMATE* members, which showed relatedness with Ethylene, Abscisic acid (ABA) and Methyl jasmonate (MeJA) responses, respectively. A total of nine *cis*-elements related to plant growth and development were also identified in some *MdMATE* gene promoter regions and the metabolism regulation element (O<sub>2</sub>-site) was found in 20 *MdMATE* genes. Various regulatory elements were identified, suggesting that the *MdMATE* genes play a crucial role in a wide range of biological processes in apple.

### Expression of *MdMATE* response to pathogen infection

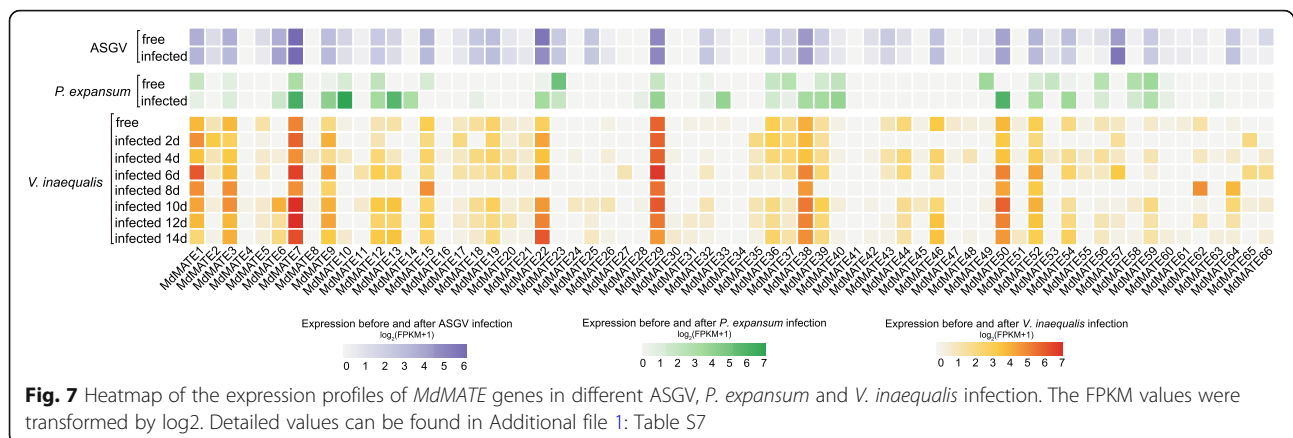
In apple and other fruits, the quality and yield can be greatly reduced in the presence of diseases [43]. An increasing body of evidence shows that the *MATE* genes are essential in conferring tolerance to abiotic and biotic stress factors [11, 21, 44]. To explore the role of *MdMATE* genes in response to disease, we analyzed the changes of transcription levels of *MdMATE* in response to *Apple stem grooving virus* (ASGV), *Penicillium expansum* (*P. expansum*) and *Venturia inaequalis* (*V. inaequalis*). These three pathogens affect the quality and yield of apples, cause a lot of economic losses and waste. In the control before and after the ASGV infection, the expression levels of *MdMATE57* showed up-regulated, while *MdMATE22* showed down-regulated (Fig. 7). After *P. expansum* infection, a total of 16 and 7 genes were up-regulated and down-regulated, respectively (Fig. 7). *MdMATE7*, *MdMATE10*, *MdMATE13* and *MdMATE50* barely expressed in *P. expansum* free, while their expression level in terms of FPKM increases to 94.76, 182.35, 65.64 and 82.61 after *P. expansum* infection, respectively (Additional file 1: Table S7). In contrast, *MdMATE23*,

*MdMATE37* and *MdMATE49* showed dramatic decrease after *P. expansum* infection. Apple scab is the most serious disease of apple worldwide in terms of the economic cost of control [45, 46]. It is caused by *V. inaequalis*, a fungus that can cause the apple disease. Scab can be found in almost all areas in which apples are grown commercially. Here we analyzed the *MdMATE* gene expression levels of different stages after *V. inaequalis* infection (Fig. 7). Some genes showed up-regulated after infecting by *V. inaequalis* such as *MdMATE2*, *MdMATE7* and *MdMATE9*, whereas some genes showed down-regulated such as *MdMATE3*, *MdMATE15* and *MdMATE46*. The expression level of *MdMATE6*, *MdMATE59*, *MdMATE62* and *MdMATE64* increased after 8 days of infection. These results indicated the *MATE* gene family in apple have a diverse function in responses to biotic stress, especially for *P. expansum* infection.

### Discussion

Recently, the genome-wide analysis of *MATE* gene family in different species has been gradually carried out, suggesting their diverse function involved in stress responses, plant growth and transmembrane transport. In this study, we carried out a comprehensive analysis of *MATE* gene family in apple. We identified 66 *MdMATE* genes from the ‘Golden Delicious’ reference genome, including two genes that have been reported. We conducted phylogenetic, gene structure, motif pattern, synteny and evolutionary analysis. Besides, the expression patterns of *MdMATE* genes in different tissues/organs and developmental stages were also explored. The *cis*-element analysis shows that *MdMATE* genes are widely involved in stress response in apples. Hence, we also compared the changes in *MdMATE* gene expression before and after three different pathogen infections.

A total of 66 *MdMATE* genes were identified in apple. The *MATE* members in apple exceed those in *Arabidopsis* (56 members, ~ 135 Mb) [10], potato (48 members, ~



674 Mb) [6], maize (49 members, ~2300 Mb) [15] and rice (45 members, ~380 Mb) [16], but family size is inferior to soybean (117 members, ~978 Mb) [17]. The number of *MATE* genes is not paralleled to the size of genome and there is no absolute correlation with genome size. However, in dicots, it was paralleled to the genome size and generally have more members than monocots. Further synteny analysis showed there are more collinear gene pairs between apple and dicots than monocots. We speculated that the expansion of *MATE* gene family mainly occurs during the evolution of dicots.

Phylogenetic analysis divides the *MdMATE* genes into four groups based on the classifications on *AtDTX* proteins (Fig. 1). In group I, the *MdMATE* lost five members, while group II, III and IV gained four, three and eight members, respectively. Group I contains 14 *MdMATE* transporters along with 19 *AtDTX* transporters. Among them, the *AtDTX1*, which is the first plant protein that has been shown to function as a detoxifying efflux carrier, can mediate the efflux of plant-derived or exogenous toxic compounds from the cytoplasm [10]. The *AtDTX14* and *AtDTX18* can export norfloxacin and prevent *Arabidopsis* colonization by *P. infestans*, respectively [47, 48]. The *AtDTX19* is expressed in root epidermal cells and necessary for protecting roots from toxic compounds in the soil [12]. Group II contain 20 *MdMATE* transporters along with 16 *AtDTX* transporters. Among them, the *AtDTX21* plays an important role in Atrazine detoxification [49]. The *AtDTX30*, *AtDTX31*, *AtDTX33* and *AtDTX35* control root hair development in *Arabidopsis* [50, 51]. Besides, the *AtDTX30* also involve in aluminum tolerance and indirectly modulated citrate exudation [52]. Group III contain 9 *MdMATE* transporters along with 6 *MATE* transporters from *Arabidopsis*. The *AtDTX41/TRANSPARENT TESTA 12 (TT12)* had been functionally characterized that can mediate anthocyanin transport in vitro [13]. Its homologs in apple, *MdMATE1* and *MdMATE2*, are vacuolar flavonoid transporter which are active in proanthocyanidins accumulating cells of apple fruit [22]. In the two subgroups of group IV, the group IV-a contain 9 and 6 *MATE* transporters from apple and *Arabidopsis*, respectively. The *AtDTX42* can facilitate AI-activated citrate exudation independently from and along with *AtALMT1* to confer a full expression of *Arabidopsis* AI tolerance [53, 54]. The *AtDTX43* plays a major role in iron and zinc homeostasis in *Arabidopsis* and transports citrate [55]. The *AtDTX47* is related to the transport of salicylic acid or its precursor [56]. While 14 *MdMATE* transporters and 9 *AtDTX* transporters in the group IV-b. The *AtDTX48* is related to multiple functions, including organ initiation, iron homeostasis and hypocotyl cell elongation [57–59]. The *AtDTX50*

can transport abscisic acid (ABA) and respond to drought conditions [18]. The function of *AtDTX51* is related to hypocotyl cell elongation same as *AtDTX48*, but additionally has the functions of regulates plant disease resistance and affects plant architecture [60, 61]. The *AtDTX51* and *AtDTX52* control senescence and iron homeostasis in plants [62, 63]. The *AtDTX54* and *AtDTX55* have a conserved function in the regulation of lateral organ initiation in plants [19]. The *AtDTX56* can repress a protein kinase that negatively regulates CO<sub>2</sub>-induced stomatal closing [64]. From the functions of these homologous *AtDTX* transporters in the subgroups of group IV, the function of group IV-a is mainly related to ion tolerance and in group IV-b, it is also related to developmental and disease resistance in addition to the functions similar as group IV-a. Overall, *Arabidopsis* *MATE* members in the same group have highly diverse functions, indicating *MATE* in apple may also have multitasking ability.

In group I, only four *AtDTX* genes (*AtDTX1*, *AtDTX14*, *AtDTX18* and *AtDTX19*) [10, 12, 47, 48] were functionally characterized but no duplicated *MdMATE* genes pairs clustered together with these *AtDTX* genes. However, in group IV, duplicated *MdMATE* genes pairs such as *MdMATE12/MdMATE40*, *MdMATE15/MdMATE44*, *MdMATE18/MdMATE36*, *MdMATE57/MdMATE32*, *MdMATE53/MdMATE10*, *MdMATE45/MdMATE16*, *MdMATE17/MdMATE35* and *MdMATE51/MdMATE63* were clustered together with *AtDTX42*, *AtDTX44*, *AtDTX46*, *AtDTX48*, *AtDTX51*, *AtDTX53*, *AtDTX54* and *AtDTX55*, respectively. Some of these *AtDTX* genes are involved in diverse mechanisms that are indispensable to plant growth and development: the *AtDTX42* is related with AI tolerance [53, 55]; the *AtDTX48*, *AtDTX51*, *AtDTX54* and *AtDTX55* all related to plant growth [19, 57–61]. The duplicated *MdMATE* gene pairs retained during the evolution process of apple may have similar functions with these neighbor *AtDTX* genes.

Different expression patterns of *MdMATE* genes were observed in various tissues/organs and developmental stages. It is worth noting that *MdMATE7* and *MdMATE36* are expressed in all samples in our study. These two genes may be important for maintenance of apple growth and development. Furthermore, expression correlation analysis reveals that *MdMATE* duplicates may follow different functional models (Additional file 1: Table S3). *MdMATE51/MdMATE63* have strong positive correlation (Spearman's  $\rho = 0.878$ ,  $P < 0.001$ ) in flower. Both their expression levels were relatively high. The corresponding gene dosage increase may be beneficial for organism [38]. *MdMATE22/MdMATE58* show strong negative correlations (Spearman's  $\rho = -0.882$ ,  $P < 0.001$ ) in fruit developmental stages. In the early stages,

*MdMATE22* expression level was higher than *MdMATE58*, but opposite in the late fruit development stage. This suggested that sub-functionalization probably occurred between the pair of genes [38, 65].

*Cis*-elements analysis further confirmed the versatility of MATE gene family and its major role in response to stress. By analyzing changes in the expression levels of *MdMATE* genes before and after infection of three pathogens that have a great impact on the quality and yield of apples, suggesting that *MdMATE* genes are involved in apple response to pathogen infection, especially for *P. expansum* infection. Notably, *MdMATE7* is involved in both infections by *P. expansum* and *V. inaequalis* (Fig. 7), and this gene is expressed in all tissue/organs. Thus, we speculated that the sustained expression of *MdMATE7* is important for apple growth and stress response. Additionally, the promoter regions of all *MdMATE* genes contain many stress response *cis*-elements and we further speculated that the other genes, which are not involved in the three pathogens infections has great potential for stress response.

## Conclusions

In this study, a total of 66 MATE (*MdMATE*) genes encoding MATE transporters were identified in the apple genome. We classified these *MdMATE* genes into four groups by phylogenetic analysis with MATE genes in *Arabidopsis*. Synteny analysis reveals that whole genome duplication and segmental duplication events played a major role in the expansion of MATE gene family in apple. *MdMATE* genes show diverse expression patterns in different tissues/organs and developmental stages. Analysis of *cis*-regulatory elements in *MdMATE* promoter regions indicates that the function of *MdMATE* gene is mainly related to stress response. Besides, the changes of gene expression levels upon different pathogen infections reveal that *MdMATE* genes are involved in biotic stress response. Our results provide insights for a more comprehensive understanding of the MATE gene family function in apple and provide valuable resources for apple disease resistance research.

## Methods

### Identification of MATE transporters in apple genome

A total of 56 MATE family members in *Arabidopsis* as previously reported [10] were download from The Arabidopsis Information Resource (TAIR) (<https://www.arabidopsis.org/>). A total of 101 apple putative MATE protein sequences were retrieved by BLASTP searches against the target apple proteome, GDDH13 v1.1 [23], using 56 *A.thaliana* MATE protein sequences as queries ( $E\text{-value} \leq 10^{-7}$ ) (Additional file 1: Table S1). To obtain more accurate MATE members in apple, manual filtering for the putative MATE protein sequences was

performed as previously described [28]. Briefly, the putative MATE protein sequences were filtered by the presence of conserved MATE domain (Pfam: PF01554) using the HMMER (<https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan>) [66], the Conserved Domain Database (CDD, <https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) [67] and the Simple Modular Architecture Research Tool (SMART, <http://smart.embl-heidelberg.de/smart/batch.pl>) [68]. All these putative sequences were assessed against the expected features of the MATE transporters in plants (e.g., containing MatE domains, 8–12 transmembrane domains and classified as MATE\_like superfamily) (Additional file 1: Table S1). Finally, a total of 66 apple MATE proteins were identified. Physical parameters such as theoretical isoelectric point (pI), molecular weight (MW) and instability index (II) were calculated by ProtParam (<https://web.expasy.org/protparam/>) [69]. The subcellular localization of the MATE proteins were predicted using WoLF PSORT [70].

### Chromosomal distribution and gene duplication analysis of *MdMATE* family

The physical location information of all 66 *MdMATE* genes were obtained from the apple genome annotation gff3 format file and visualize through Circlize package [71]. WGD/segmental and tandem duplication events were detected by MCScanX [72] with default parameters. Tandem clusters were defined as previously report: genes in a cluster need to be on the same chromosome and not more than one gene apart [73]. To exhibit the synteny relationship of the orthologous MATE genes in apple and other four species, MATE gene family members in *Arabidopsis*, soybean, maize and rice were obtained as the previous report [10, 15–17] and sequence files were download from Phytozome database (v12, <https://phytozome.jgi.doe.gov/pz/portal.html>) [74]. Species tree were obtained from TimeTree (<http://www.timetree.org>) [75]. Then MCScanX pipeline [33] was used to construct syntenic maps and visualization. Ka/Ks values between homologous were calculated by KaKs\_Calculator 2.0 [76].

### Phylogenetic and gene structural analysis

All the 122 MATE protein sequences in apple and *Arabidopsis* were used to perform phylogenetic and structural analysis. Multiple sequence alignments were conducted by ClustalW in MEGA X with default parameters and ProteinWeightMatrix = BLOSUM [77]. The alignment result was then used to construct a phylogenetic tree based on the neighbor-joining (NJ) method of MEGA X, with the following setups: Equal input model and partial deletion (60%). For the maximum likelihood (ML) tree, JTT + F + G was used as the best model for ML tree constructing which calculated by ProtTest 3.4.2

[78], the ML tree was constructed by MEGA X. Both NJ and ML trees were conducted 1000 bootstrap replications.

Gene structure analysis was performed using the Gene Structure Display Server (GSDS) [79] with default setting. Motifs in MATE proteins were identified using MEME Suite [25] (version 5.1.1, <http://meme-suite.org/index.html>) with default setting: site distribution is zero or one occurrence per sequence (zoops), motif width is 5 to 50 and the maximum number of motifs was set at 12. We extract the 1.5 kb upstream region of *MdMATE* genes and upload to PlantCARE [80] database to detect *cis*-regulatory elements in the promoter regions.

### Expression pattern of *MdMATE* genes in different tissues and pathogen infection

We obtain a total of 48 expression profiles of 66 *MdMATE* genes from Apple Multi-Dimensional Omics Database (AppleMDO) [35]. These profiles including 36 tissues/organs and different development stage (central seed, lateral seed, stem, leaf, flower, petal, stigma, style, ovary, anther, filament, sepal, receptacle, pollen, four dormant bud stage, break bud, fourteen fruit developmental stages from 1 week after full-bloom (WAF1) to harvest (WAF20), ripe fruit peel and flesh and control data of three pathogens infection. In the database, all the RNA-seq data were already quality controlled and FPKM values can be extracted from the database. Expression heatmaps were made by Pheatmap package with pheatmap2 function [81]. Spearman correlation was conducted using the cor() function in R-4.0.2 (<https://www.r-project.org/>).

### Abbreviations

ABA: Transport abscisic acid; ABC: ATP-binding cassette superfamily; ASGV: Apple stem grooving virus; FPKM: Fragments Per Kilobase per Million; MATE: Multidrug and toxic compound extrusion; MFS: Major facilitator superfamily; MW: Molecular weight; NJ: Neighbor-joining; pI: Theoretical isoelectric point; RND: Resistance/nodulation/division family; SMR: Small gene multidrug resistance family; WGD: Whole genome duplication

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-021-07943-1>.

**Additional file 1 Table S1.** Manual filtering of 101 putative MATE proteins. **Table S2.** Details of the 66 *MdMATE* genes in apple. **Table S3.** Duplicated gene pairs of *MdMATE* genes and expression pattern correlations. **Table S4.** Collinear gene pairs between *MdMATE* genes and MATE genes in other species. **Table S5.** FPKM of *MdMATE* genes in different tissues/organs and developmental stages. **Table S6.** *Cis*-elements of *MdMATE* genes. **Table S7.** FPKM of *MdMATE* genes with three pathogen infections.

**Additional file 2 Fig. S1.** The unrooted neighbor-joining phylogenetic tree of MATE family members in apple and Arabidopsis. The different colors indicate different groups (Group I in blue, Group II in orange, Group III in green and Group IV in pink). 'MdMATE' represents MATE members from apple, 'AtDTX' represents MATE members from Arabidopsis. Numbers on the nodes are bootstrap values in percentage (1000

replicates). **Fig. S2.** The conserved motifs among *MdMATE* proteins. **Fig. S3.** Spearman correlation of gene expression pattern in flower, bud and fruit. Positive correlations are displayed in red and negative correlations in blue color. Size and color intensity of the square are proportional to the Spearman's  $\rho$ . Red and yellow lines on the top of correlation heatmap indicate segmental and tandem duplicated pairs, respectively.

### Acknowledgements

The authors thank to lab members for assistance.

### Authors' contributions

Li.L. and W.Z. conceptualization this study. W.Z., Liao.L. and J.X. perform bioinformatic analyses. Liao.L. and Y.H. helped with results interpretation. W.Z. wrote the initial manuscript. Li.L. and Y.H. critically reviewed and edited the article. All authors have read and agreed to the published version of the manuscript. The authors read and approved the final manuscript

### Funding

The research was funded by Huazhong Agricultural University Scientific & Technological Self-innovation Foundation to Li.L. and supported by Hubei Hongshan Laboratory to Li.L. The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

### Availability of data and materials

The reference genome were GDDH13 Version 1.1 and obtained from The Apple Genome and Epigenome (<https://iris.angers.inra.fr/gddh13/the-apple-genome-downloads.html>). All data analyzed during this study are included in this article and its Additional files. *MdMATE* gene names and details can be found in Additional file 1: Table S1 and S2. The Ka/Ks and spearman's correlation coefficient of each gene pairs can be found in Additional file 1: Table S3 and Table S4. The FPKM value of *MdMATE* genes in different tissues/organs and pathogen infections can be found in Additional file 1: Table S5 and Table S7. The *cis*-elements annotations of *MdMATE* genes can be found in Additional file 1: Table S5.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Hubei Key Laboratory of Agricultural Bioinformatics, College of Informatics, Huazhong Agricultural University, Wuhan 430070, People's Republic of China. <sup>2</sup>CAS Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, The Innovative Academy of Seed Design, Chinese Academy of Sciences, Wuhan 430074, People's Republic of China. <sup>3</sup>Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan 430070, People's Republic of China.

Received: 30 March 2021 Accepted: 12 August 2021

Published online: 30 August 2021

### References

- Omote H, Hiasa M, Matsumoto T, Otsuka M, Moriyama Y. The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. *Trends Pharmacol Sci.* 2006;27(11):587–93. <https://doi.org/10.1016/j.tips.2006.09.001>.
- Remy E, Duque P. Beyond cellular detoxification: a plethora of physiological roles for MDR transporter homologs in plants. *Front Physiol.* 2014;5:201.
- Putman M, van Veen HW, Konings WN. Molecular properties of bacterial multidrug transporters. *Microbiol Mol Biol Rev.* 2000;64(4):672–93. <https://doi.org/10.1128/MMBR.64.4.672-693.2000>.

4. Wang R, Zhao P, Kong N, Lu R, Pei Y, Huang C, Ma H, Chen Q. Genome-Wide Identification and Characterization of the Potato bHLH Transcription Factor Family. *Genes* (Basel). 2018;9(1).
5. Saier MH Jr. Genome archeology leading to the characterization and classification of transport proteins. *Curr Opin Microbiol*. 1999;2(5):555–61. [https://doi.org/10.1016/S1369-5274\(99\)00016-8](https://doi.org/10.1016/S1369-5274(99)00016-8).
6. Li Y, He H, He LF. Genome-wide analysis of the MATE gene family in potato. *Mol Biol Rep*. 2019;46(1):403–14. <https://doi.org/10.1007/s11033-018-4487-y>.
7. Morita Y, Kodama K, Shiota S, Mine T, Kataoka A, Mizushima T, et al. NorM, a putative multidrug efflux protein, of *Vibrio parahaemolyticus* and its homolog in *Escherichia coli*. *Antimicrob Agents Chemother*. 1998;42(7):1778–82. <https://doi.org/10.1128/AAC.42.7.1778>.
8. Pao SS, Paulsen IT, Saier MH. Major facilitator superfamily. *Microbiol Mol Biol Rev*. 1998;62(1):1–34. <https://doi.org/10.1128/MMBR.62.1.1-34.1998>.
9. Brown MH, Paulsen IT, Skurray RA. The multidrug efflux protein NorM is a prototype of a new family of transporters. *Mol Microbiol*. 1999;31(1):394–5. <https://doi.org/10.1046/j.1365-2958.1999.01162.x>.
10. Li L, He Z, Pandey GK, Tsuchiya T, Luan S. Functional cloning and characterization of a plant efflux carrier for multidrug and heavy metal detoxification. *J Biol Chem*. 2002;277(7):5360–8. <https://doi.org/10.1074/jbc.M108777200>.
11. Upadhyay N, Kar D, Deepak Mahajan B, Nanda S, Rahiman R, Panchakshari N, et al. The multitasking abilities of MATE transporters in plants. *J Exp Bot*. 2019;70(18):4643–56. <https://doi.org/10.1093/jxb/erz246>.
12. Diener AC, Gaxiola RA, Fink GR. Arabidopsis ALF5, a multidrug efflux transporter gene family member, confers resistance to toxins. *Plant Cell*. 2001;13(7):1625–38. <https://doi.org/10.1105/TPC.010035>.
13. Marinova K, Pourcel L, Weder B, Schwarz M, Barron D, Routaboul JM, et al. The Arabidopsis MATE transporter TT12 acts as a vacuolar flavonoid/H<sup>+</sup>-antiporter active in proanthocyanidin-accumulating cells of the seed coat. *Plant Cell*. 2007;19(6):2023–38. <https://doi.org/10.1105/tpc.106.046029>.
14. Santos ALD, Chaves-Silva S, Yang L, Maia LGS, Chalfun-Junior A, Sinharoy S, et al. Global analysis of the MATE gene family of metabolite transporters in tomato. *BMC Plant Biol*. 2017;17(1):185. <https://doi.org/10.1186/s12870-017-1115-2>.
15. Zhu H, Wu J, Jiang Y, Jin J, Zhou W, Wang Y, et al. Genomewide analysis of MATE-type gene family in maize reveals microsynteny and their expression patterns under aluminum treatment. *J Genet*. 2016;95(3):691–704. <https://doi.org/10.1007/s12041-016-0686-2>.
16. Wang L, Bei X, Gao J, Li Y, Yan Y, Hu Y. The similar and different evolutionary trends of MATE family occurred between rice and *Arabidopsis thaliana*. *BMC Plant Biol*. 2016;16(1):207. <https://doi.org/10.1186/s12870-016-0895-0>.
17. Liu J, Li Y, Wang W, Gai J, Li Y. Genome-wide analysis of MATE transporters and expression patterns of a subgroup of MATE genes in response to aluminum toxicity in soybean. *BMC Genomics*. 2016;17(1):223. <https://doi.org/10.1186/s12864-016-2559-8>.
18. Zhang H, Zhu H, Pan Y, Yu Y, Luan S, Li L. A DTX/MATE-type transporter facilitates abscisic acid efflux and modulates ABA sensitivity and drought tolerance in *Arabidopsis*. *Mol Plant*. 2014;7(10):1522–32. <https://doi.org/10.1093/mp/ssu063>.
19. Suzuki M, Sato Y, Wu S, Kang BH, McCarty DR. Conserved functions of the MATE transporter BIG EMBRYO1 in regulation of lateral organ size and initiation rate. *Plant Cell*. 2015;27(8):2288–300. <https://doi.org/10.1105/tpc.15.00290>.
20. Sheikh MA, Mushtaq K, Mir JI, Amin M, Nabi SU. Introgression of scab resistance gene Vf (Rvi6) in commercially grown susceptible cultivar Fuji Azitec of apple (*Malus domestica*) using marker assisted selection. *Res J Biotechnol Vol*. 2020;15:9.
21. Tiwari M, Sharma D, Singh M, Tripathi RD, Trivedi PK. Expression of OsMATE1 and OsMATE2 alters development, stress responses and pathogen susceptibility in *Arabidopsis*. *Sci Rep*. 2014;4:3964.
22. Frank S, Keck M, Sagasser M, Niehaus K, Weisshaar B, Stracke R. Two differentially expressed MATE factor genes from apple complement the *Arabidopsis* transparent testa12 mutant. *Plant Biol*. 2011;13(1):42–50. <https://doi.org/10.1111/j.1438-8677.2010.00350.x>.
23. Daccord N, Celton JM, Linsmith G, Becker C, Choise N, Schijlen E, et al. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nat Genet*. 2017;49(7):1099–106. <https://doi.org/10.1038/ng.3886>.
24. Li N, Meng H, Xing H, Liang L, Zhao X, Luo K. Genome-wide analysis of MATE transporters and molecular characterization of aluminum resistance in *Populus*. *J Exp Bot*. 2017;68(20):5669–83. <https://doi.org/10.1093/jxb/erx370>.
25. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*. 2009;37(Web Server issue):W202–8.
26. Yamasaki K, Kigawa T, Seki M, Shinozaki K, Yokoyama S. DNA-binding domains of plant-specific transcription factors: structure, function, and evolution. *Trends Plant Sci*. 2013;18(5):267–76. <https://doi.org/10.1016/j.tplants.2012.09.001>.
27. Xu L, Shen ZL, Chen W, Si GY, Meng Y, Guo N, et al. Phylogenetic analysis of upland cotton MATE gene family reveals a conserved subfamily involved in transport of proanthocyanidins. *Mol Biol Rep*. 2019;46(1):161–75. <https://doi.org/10.1007/s11033-018-4457-4>.
28. Ali E, Saand MA, Khan AR, Shah JM, Feng S, Ming C, et al. Genome-wide identification and expression analysis of detoxification efflux carriers (DTX) genes family under abiotic stresses in flax. *Physiol Plant*. 2020;171(4):483–501. <https://doi.org/10.1111/pp.13105>.
29. See DR, Brooks S, Nelson JC, Brown-Guedira G, Friebe B, Gill BS. Gene evolution at the ends of wheat chromosomes. *Proc Natl Acad Sci U S A*. 2006;103(11):4162–7. <https://doi.org/10.1073/pnas.0508942102>.
30. Moore RC, Purugganan MD. The early stages of duplicate gene evolution. *Proc Natl Acad Sci U S A*. 2003;100(26):15682–7. <https://doi.org/10.1073/pnas.2535513100>.
31. Xie T, Chen C, Li C, Liu J, Liu C, He Y. Genome-wide investigation of WRKY gene family in pineapple: evolution and expression profiles during development and stress. *BMC Genomics*. 2018;19(1):490. <https://doi.org/10.1186/s12864-018-4880-x>.
32. Cannon SB, Mitra A, Baumgarten A, Young ND, May G. The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol*. 2004;4(1):10. <https://doi.org/10.1186/1471-2229-4-10>.
33. Tang H, Bowers JE, Wang X, Ming R, Alam M, Paterson AH. Synteny and collinearity in plant genomes. *Science*. 2008;320(5875):486–8. <https://doi.org/10.1126/science.11553917>.
34. Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, et al. Genome sequence of the palaeopolyploid soybean. *Nature*. 2010;463(7278):178–83. <https://doi.org/10.1038/nature08670>.
35. Da L, Liu Y, Yang J, Tian T, She J, Ma X, et al. AppleMDO: a multi-dimensional omics database for apple co-expression networks and chromatin states. *Front Plant Sci*. 2019;10:1333. <https://doi.org/10.3389/fpls.2019.01333>.
36. Ohno S: Evolution by gene duplication: Springer Science & Business Media; 2013.
37. Li W-H, Yang J, Gu X. Expression divergence between duplicate genes. *Trends Genet*. 2005;21(11):602–7. <https://doi.org/10.1016/j.tig.2005.08.006>.
38. Panchy N, Lehti-Shiu M, Shiu SH. Evolution of gene duplication in plants. *Plant Physiol*. 2016;171(4):2294–316. <https://doi.org/10.1104/pp.16.00523>.
39. Schilling S, Kennedy A, Pan S, Jermin LS, Melzer R. Genome-wide analysis of MIK-type MADS-box genes in wheat: pervasive duplications, functional conservation and putative neofunctionalization. *New Phytol*. 2020;225(1):511–29. <https://doi.org/10.1111/nph.16122>.
40. Hughes TE, Langdale JA, Kelly S. The impact of widespread regulatory neofunctionalization on homeolog gene evolution following whole-genome duplication in maize. *Genome Res*. 2014;24(8):1348–55. <https://doi.org/10.1101/gr.172684.114>.
41. Buggs RJ, Elliott NM, Zhang L, Koh J, Viccini LF, Soltis DE, et al. Tissue-specific silencing of homeologs in natural populations of the recent allopolyploid *Tragopogon mirus*. *New Phytol*. 2010;186(1):175–83. <https://doi.org/10.1111/j.1469-8137.2010.03205.x>.
42. Wittkopp PJ, Kalay G. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet*. 2011;13(1):59–69. <https://doi.org/10.1038/nrg3095>.
43. Dubej SR, Jalal AS. Apple disease classification using color, texture and shape features from images. *SIVIP*. 2016;10(5):819–26. <https://doi.org/10.1007/s11760-015-0821-1>.
44. Takanashi K, Shitan N, Yazaki K. The multidrug and toxic compound extrusion (MATE) family in plants. *Plant Biotechnol*. 2014;31(5):417–30. <https://doi.org/10.5511/plantbiotechnology.14.0904a>.
45. Carisse O, Bernier J. Effect of environmental factors on growth, pycnidial production and spore germination of *Microsphaeropsis* isolates with biocontrol potential against apple scab. *Mycol Res*. 2002;106(12):1455–62. <https://doi.org/10.1017/S0953756202006858>.

46. Bowen JK, Mesarich CH, Bus VG, Beresford RM, Plummer KM, Templeton MD. *Venturia inaequalis*: the causal agent of apple scab. *Mol Plant Pathol*. 2011;12(2):105–22. <https://doi.org/10.1111/j.1364-3703.2010.00656.x>.
47. Miyauchi H, Moriyama S, Kusakizako T, Kumazaki K, Nakane T, Yamashita K, et al. Structural basis for xenobiotic extrusion by eukaryotic MATE transporter. *Nat Commun*. 2017;8(1):1633. <https://doi.org/10.1038/s41467-017-01541-0>.
48. Dobritsch M, Lubken T, Eschen-Lippold L, Gorzalka K, Blum E, Matern A, et al. MATE transporter-dependent export of Hydroxycinnamic acid amides. *Plant Cell*. 2016;28(2):583–96. <https://doi.org/10.1105/tpc.15.00706>.
49. Ramel F, Sulmon C, Cabello-Hurtado F, Taconnat L, Martin-Magniette ML, Renou JP, et al. Genome-wide interacting effects of sucrose and herbicide-mediated stress in *Arabidopsis thaliana*: novel insights into atrazine toxicity and sucrose-induced tolerance. *BMC Genomics*. 2007;8(1):450. <https://doi.org/10.1186/1471-2164-8-450>.
50. Won SK, Lee YJ, Lee HY, Heo YK, Cho M, Cho HT. Cis-element- and transcriptome-based screening of root hair-specific genes and their functional characterization in *Arabidopsis*. *Plant Physiol*. 2009;150(3):1459–73. <https://doi.org/10.1104/pp.109.140905>.
51. Zhang H, Zhao FG, Tang RJ, Yu Y, Song J, Wang Y, et al. Two tonoplast MATE proteins function as turgor-regulating chloride channels in *Arabidopsis*. *Proc Natl Acad Sci U S A*. 2017;114(10):E2036–45. <https://doi.org/10.1073/pnas.1616203114>.
52. Upadhyay N, Kar D, Datta S. A multidrug and toxic compound extrusion (MATE) transporter modulates auxin levels in root to regulate root development and promotes aluminium tolerance. *Plant Cell Environ*. 2020;43(3):745–59. <https://doi.org/10.1111/pce.13658>.
53. Liu J, Magalhaes JV, Shaff J, Kochian LV. Aluminum-activated citrate and malate transporters from the MATE and ALMT families function independently to confer *Arabidopsis* aluminum tolerance. *Plant J*. 2009;57(3):389–99. <https://doi.org/10.1111/j.1365-313X.2008.03696.x>.
54. Liu J, Luo X, Shaff J, Liang C, Jia X, Li Z, et al. A promoter-swap strategy between the AtALMT and AtMATE genes increased *Arabidopsis* aluminum resistance and improved carbon-use efficiency for aluminum resistance. *Plant J*. 2012;71(2):327–37. <https://doi.org/10.1111/j.1365-313X.2012.04994.x>.
55. Scheepers M, Spielmann J, Boulanger M, Carnol M, Bosman B, De Pauw E, et al. Intertwined metal homeostasis, oxidative and biotic stress responses in the *Arabidopsis* frd3 mutant. *Plant J*. 2020;102(1):34–52. <https://doi.org/10.1111/tpj.14610>.
56. Nawrath C, Heck S, Parinthewong N, Metraux JP. EDS5, an essential component of salicylic acid-dependent signaling for disease resistance in *Arabidopsis*, is a member of the MATE transporter family. *Plant Cell*. 2002;14(1):275–86. <https://doi.org/10.1105/tpc.010376>.
57. Burko Y, Geva Y, Refael-Cohen A, Shleizer-Burko S, Shani E, Berger Y, et al. From organelle to organ: ZRIZI MATE-type transporter is an organelle transporter that enhances organ initiation. *Plant Cell Physiol*. 2011;52(3):518–27. <https://doi.org/10.1093/pcp/pcr007>.
58. Seo PJ, Park J, Park MJ, Kim YS, Kim SG, Jung JH, et al. A Golgi-localized MATE transporter mediates iron homeostasis under osmotic stress in *Arabidopsis*. *Biochem J*. 2012;442(3):551–61. <https://doi.org/10.1042/BJ20111311>.
59. Wang R, Liu X, Liang S, Ge Q, Li Y, Shao J, et al. A subgroup of MATE transporter genes regulates hypocotyl cell elongation in *Arabidopsis*. *J Exp Bot*. 2015;66(20):6327–43. <https://doi.org/10.1093/jxb/erv344>.
60. Li R, Li J, Li S, Qin G, Novak O, Pencik A, et al. ADP1 affects plant architecture by regulating local auxin biosynthesis. *PLoS Genet*. 2014;10(1):e1003954. <https://doi.org/10.1371/journal.pgen.1003954>.
61. Sun X, Gilroy EM, Chini A, Nurnberg PL, Hein I, Lacomme C, et al. ADS1 encodes a MATE-transporter that negatively regulates plant disease resistance. *New Phytol*. 2011;192(2):471–82. <https://doi.org/10.1111/j.1469-8137.2011.03820.x>.
62. Jia M, Liu X, Xue H, Wu Y, Shi L, Wang R, et al. Noncanonical ATG8-ABS3 interaction controls senescence in plants. *Nat Plants*. 2019;5(2):212–24. <https://doi.org/10.1038/s41477-018-0348-x>.
63. Wang Z, Qian C, Guo X, Liu E, Mao K, Mu C, et al. ELS1, a novel MATE transporter related to leaf senescence and iron homeostasis in *Arabidopsis thaliana*. *Biochem Biophys Res Commun*. 2016;476(4):319–25. <https://doi.org/10.1016/j.bbrc.2016.05.121>.
64. Tian W, Hou C, Ren Z, Pan Y, Jia J, Zhang H, et al. A molecular pathway for CO(2) response in *Arabidopsis* guard cells. *Nat Commun*. 2015;6(1):6057. <https://doi.org/10.1038/ncomms7057>.
65. Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*. 1999;151(4):1531–45. <https://doi.org/10.1093/genetics/151.4.1531>.
66. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam protein families database in 2019. *Nucleic Acids Res*. 2019;47(D1):D427–32. <https://doi.org/10.1093/nar/gky995>.
67. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res*. 2020;48(D1):D265–8. <https://doi.org/10.1093/nar/gkz291>.
68. Letunic I, Doerks T, Bork P. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res*. 2015;43(Database issue):D257–60. <https://doi.org/10.1093/nar/gku949>.
69. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A: Protein identification and analysis tools on the ExPASy server. In: The proteomics protocols handbook. Springer; 2005: 571–607.
70. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, et al. WoLF PSORT: protein localization predictor. *Nucleic Acids Res*. 2007;35(Web Server issue):W585–7.
71. Gu Z, Gu L, Eils R, Schlesner M, Brors B. Circize implements and enhances circular visualization in R. *Bioinformatics*. 2014;30(19):2811–2. <https://doi.org/10.1093/bioinformatics/btu393>.
72. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res*. 2012;40(7):e49. <https://doi.org/10.1093/nar/gkr1293>.
73. Zhu Y, Wu N, Song W, Yin G, Qin Y, Yan Y, et al. Soybean (Glycine max) expansin gene superfamily origins: segmental and tandem duplication events followed by divergent selection among subfamilies. *BMC Plant Biol*. 2014;14(1):93. <https://doi.org/10.1186/1471-2229-14-93>.
74. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res*. 2012;40(Database issue):D1178–86. <https://doi.org/10.1093/nar/gkr944>.
75. Kumar S, Stecher G, Suleski M, Hedges SB. TimeTree: a resource for timelines, timetrees, and divergence times. *Mol Biol Evol*. 2017;34(7):1812–9. <https://doi.org/10.1093/molbev/msx116>.
76. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. KaKs\_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics, proteomics & bioinformatics*. 2010;8(1):77–80. [https://doi.org/10.1016/S1672-0229\(10\)60008-3](https://doi.org/10.1016/S1672-0229(10)60008-3).
77. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;35(6):1547–9. <https://doi.org/10.1093/molbev/msy096>.
78. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics*. 2011;27(8):1164–5. <https://doi.org/10.1093/bioinformatics/btr088>.
79. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*. 2015;31(8):1296–7. <https://doi.org/10.1093/bioinformatics/btu817>.
80. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res*. 2002;30(1):325–7. <https://doi.org/10.1093/nar/30.1.325>.
81. Kolde R. Pheatmap: pretty heatmaps. R package version 2012, 1(2).

## Publisher's Note

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

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

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

**At BMC, research is always in progress.**

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

