



Genome-wide evolutionary characterization and expression analyses of major latex protein (*MLP*) family genes in *Vitis vinifera*

Ningbo Zhang^{1,2,3} · Ruimin Li^{1,2,3} · Wei Shen^{1,2,3} · Shuzhen Jiao^{1,2,3} · Junxiang Zhang^{1,2,3,4} · Weirong Xu^{1,2,3}

Received: 21 March 2018 / Accepted: 19 April 2018 / Published online: 27 April 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

The major latex protein/ripening-related protein (MLP/RRP) subfamily is known to be involved in a wide range of biological processes of plant development and various stress responses. However, the biological function of MLP/RRP proteins is still far from being clear and identification of them may provide important clues for understanding their roles. Here, we report a genome-wide evolutionary characterization and gene expression analysis of the *MLP* family in European *Vitis* species. A total of 14 members, was found in the grape genome, all of which are located on chromosome 1, where are predominantly arranged in tandem clusters. We have noticed, most surprisingly, promoter-sharing by several non-identical but highly similar gene members to a greater extent than expected by chance. Synteny analysis between the grape and *Arabidopsis thaliana* genomes suggested that 3 grape *MLP* genes arose before the divergence of the two species. Phylogenetic analysis provided further insights into the evolutionary relationship between the genes, as well as their putative functions, and tissue-specific expression analysis suggested distinct biological roles for different members. Our expression data suggested a couple of candidate genes involved in abiotic stresses and phytohormone responses. The present work provides new insight into the evolution and regulation of *Vitis MLP* genes, which represent targets for future studies and inclusion in tolerance-related molecular breeding programs.

Keywords *Vitis*-major latex protein and ripening-related protein (MLP/RRP) · Phylogenetic analysis · Environmental stress · Expression analysis

Communicated by S. Hohmann.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00438-018-1440-7>) contains supplementary material, which is available to authorized users.

✉ Junxiang Zhang
zhangjunxiang@126.com

✉ Weirong Xu
xuwr@nxu.edu.cn

¹ School of Agronomy, Ningxia University, Yinchuan 750021, Ningxia, People's Republic of China

² Ningxia Engineering and Technology Research Center of Grape and Wine, Ningxia University, Yinchuan 750021, Ningxia, People's Republic of China

³ Engineering Research Center of Grape and Wine, Ministry of Education, Ningxia University, Yinchuan 750021, Ningxia, People's Republic of China

⁴ School of Vine and Wine, Ningxia University, Yinchuan 750021, Ningxia, People's Republic of China

Introduction

Grapevine (*Vitis* spp.) represents an economically important fruit crop worldwide, and is widely used for the production of wines, table grapes, raisins, juices. Recent studies have shown that the consumption of grapes and its products has numerous nutritional values and health benefits for humans, due to the presence of unique mix of polyphenolic compounds (Wu et al. 2001; Biais et al. 2017). However, grapevines are constantly exposed to a variety of biotic and abiotic stresses that reduce their fitness and performance, leading to yield loss largely and thereby causing huge amount of economic loss as well. The mechanisms of perception of such extracellular stimuli, the molecular and biochemical events that promote survival under extreme conditions, and the identity of the gene networks that control physiological responses to stress in grapevine are not well understood. Therefore, it is urgent need to understand the molecular mechanism and genes involved to minimize various stresses for mitigating *Vitis* crop losses.

Example of genes associated with stress responses are those encoding major latex protein/ripening-related proteins (MLP/RRP), the second largest subfamily of the Bet v 1 family, which have been identified in lactiferous plants, such as plantain (*Musa × paradisiaca*) (Nessler et al. 1985), and opium poppy (Osmark et al. 1998). In grapevine, little is known about the roles of *MLP* members. Full, or draft sequences of plant genomes facilitating global identification of *MLP* genes, and after a global sequence alignment, one study found 60 plant sequences belonging to the MLP/RRP subfamily. These comprised 3 sequences from 3 species of monocotyledons and 57 sequences from 22 species of dicotyledons, with 31 members in *A. thaliana* and at least 4 in *Vitis vinifera* (Radauer et al. 2008). The sequences showed distant sequence similarity and were reported to share a similar Bet v 1 fold and to function in development, stress responses and defense (Radauer et al. 2008; Lytle et al. 2009; Lee et al. 2011).

MLPs were shown to be distributed in developing and mature laticifers of opium poppy (*Papaver somniferum*) and to be excreted to injured surfaces (Nessler et al. 1985, 1990; Griffing and Nessler 1989; Nessler and Burnett 1992). Since the first isolation of a *MLP* from opium poppy latex (Nessler et al. 1985), members have been reported that are associated with responses to plant hormones (Ruperti et al. 2002; Wang et al. 2015; Yang et al. 2015), alkaloid biosynthesis (Facchini and Park 2003; Weid et al. 2004; Kutchan 2005), oxidative stress (Stanley Kim et al. 2005), salt stress (Chen and Dai 2010), plant development (Guo et al. 2011a; Litholdo et al. 2016; Nawrot et al. 2017), the transport of persistent organic pollutants (Inui et al. 2013) and defense against pathogen infection (Yang et al. 2015). The expression of *MLP* genes has also been closely with fruit ripening in wild strawberry (*Fragaria vesca*) (Nam et al. 1999), peach (*Prunus persica*) (Ruperti et al. 2002), kiwifruit (*Actinidia deliciosa*) (Chruszcz et al. 2013) and *Cucurbita pepo* ssp. *ovifera* cultivars Patty Green and Gold Rush (Inui et al. 2013). Additionally, it was reported that ectopic overexpression of the cotton (*Gossypium hirsutum*) *GhMLP28* gene in *A. thaliana* enhanced tolerance to salt stress (Chen and Dai 2010), while overexpression in *Nicotiana tabacum* enhanced resistance to *Verticillium dahliae* infection (Qu et al. 2005; Yang et al. 2015). It has also been shown that overexpression of *A. thaliana AtMLP43* improves drought tolerance by mediating abscisic acid (ABA) signal transduction (Wang et al. 2015).

An increasing number of reports describing the function of *MLPs* in plants has revealed that they may play important roles in response to biotic and abiotic stresses. However, genome-wide identification and functional analysis of the *MLP* family in *Vitis vinifera* have not yet been determined. Since the release of the grape genome sequence (Jaillon et al. 2007), a developed multi-omics data analytics platform

on grapevine have been used for gene functional analysis (Fasoli et al. 2012). This provides an opportunity to identify and characterize the grapevine *MLP* gene family. For example, microarray analysis of 54 tissues and organs at different developmental stages provided insight into the transcriptional programs in *V. vinifera* cv. Corvina (Fasoli et al. 2012). This analysis showed that most *MLP* genes are expressed through berry development (Fasoli et al. 2012). Furthermore, two *MLP28* proteins (Q9M4G9 and F6HFH2), from *V. rotundifolia* Michx. Muscadine, were shown to be up-regulated during berry ripening by iTRAQ-based quantitative proteomics, while their expression was down-regulated in *V. vinifera* cv. Muscat Hamburg (Kambiranda et al. 2013). It therefore, appears that *MLPs* are associated with fruit development and function in the defense against environmental stress.

As one of the famous red wine grape, *V. vinifera* cv. Cabernet Sauvignon is the most widely cultivated grape variety in the world. The recently released assemble sequence of the Cabernet Sauvignon genome (Chin et al. 2016) offers an opportunity to identify and analyze *MLP* gene families in this species. In the current study, we totally identified 14 *MLP* genes from *V. vinifera*. The comprehensive analysis of gene structure, chromosomal location, classification and evolution, and *cis*-acting regulatory elements located in their upstream regions were further investigated. The tissue-specific expression profiles and their responsiveness to abiotic stress and phytohormone treatments were evaluated. We present a genome-wide evolutionary characterization and expression analysis of *MLP* genes in *V. vinifera*, an important step for further investigation into their function.

Materials and methods

Identification and classification of putative *Vitis* MLP-like proteins

The hidden Markov model (HMM) conserved sequence from the Pfam database (<http://pfam.janelia.org/>) was used to search against the Refseq database to find candidate *MLP* sequences on the HMMER website (<http://hmmer.janelia.org/>). All candidate *MLP* sequences from 13 plant genomes, consisting of two monocots (*Brachypodium distachyon*, and *Setaria italica*) and 11 dicots (*A. thaliana*, *Cicer arietinum*, *Citrus sinensis*, *Cucumis sativus*, *Fragaria vesca*, *Glycine max*, *Medicago truncatula*, *Ricinus communis*, *Solanum lycopersicum*, *Solanum tuberosum* and *V. vinifera*) were analyzed using the NCBI Conserved Domains Search website tool (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) to ensure that they contained the Bet v 1 allergen domain. Finally, an online NCBI tool (<http://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwcm>

t.cgi) was used to generate a species tree. The information of genome size for each species was obtained from Phytozome v12.1 (<https://phytozome.jgi.doe.gov/pz/portal.html>).

Phylogenetic tree, chromosomal localization, and duplication events of *V. vinifera* MLP genes

Sequence alignments were performed using Clustal W (Thompson et al. 2002), and the aligned sequences were used to construct an unrooted phylogenetic trees with the neighbor-joining (NJ) method (Saitou and Nei 1987) using MEGA 6.0 software (Tamura et al. 2011), with the complete-deletion option. The confidence level of monophyletic groups was estimated using a bootstrap analysis of 1000 replicates. The grape MLP gene duplication events, as well as those syntenic events between *Vitis* and *A. thaliana* MLP genes, were identified using the Plant Genome Duplication Database (<http://chibba.agtec.uga.edu/duplication/>). The chromosomal distribution of grape MLP genes was analyzed on the Grape Genome Browser (12×) (<http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/>) and Grape Genome CRIBI database (<http://genomes.cribi.unipd.it/grape/bblast/bblast.php>).

Analysis of VvMLP exon–intron structures and conserved motifs

The exon–intron organization of the *VvMLP* genes was determined by comparing predicted coding sequences with their corresponding genomic sequences, using GSDS software (<http://gsds.cbi.pku.edu.cn>) with the default settings. Conserved motifs in the predicted *VvMLP* protein sequences were identified using MEME 4.12.0 (http://meme.sdsc.edu/meme4_4_0/cgi-bin/meme.cgi) as previously described (Verma et al. 2010). The optimized parameters of MEME were used as follows: number of repetitions, any; maximum number of motifs, 10; and the optimum width of each motif, between 6 and 50 residues.

Analysis of cis-acting regulatory elements in the upstream regions of *VvMLPs* in Cabernet Sauvignon

The 1.0 kb sequences upstream from the start codons of the *VvMLP* genes, extracted from assembled genome sequence of *Vitis vinifera* cv. Cabernet Sauvignon (Chin et al. 2016) based on their scaffold location using a self-written perl script, were defined as putative promoter regions. Identification of the potential cis-acting regulatory elements was conducted using the MatInspector program of Genomatix software suite v3.10 with the standard (0.75) core similarity and the optimized matrix similarity (<http://www.genomatix.de/matinspector.html>). The identified cis-acting regulatory

elements were drawn using Genepalette software (Rebeiz and Posakony 2004).

Microarray data analysis

The microarray data related to the different organs at various developmental stages in *V. vinifera* cv. Corvina were retrieved from Grape eFP Browser microarray database (http://bar.utoronto.ca/efp_grape/cgi-bin/efpWeb.cgi). The expression data obtained were normalized based on the mean expression value of each gene in all tissues/organs analyzed using Z score transformation (Cheadle et al. 2003). The normalized dataset was calculated by Pearson correlation analysis and clustered by hierarchical average-linkage assay with MultiExperiment Viewer (MeV v4.8, <http://www.tm4.org/mev/>).

Plant materials and treatments

Two-year-old *V. vinifera* cv. Cabernet Sauvignon potted plants developed from stem cuttings were grown in the greenhouse at the Ningxia University (Yinchuan, China) under a natural photoperiod. Uniform cutting seedlings of 50–60 cm height were used for all experiments. Before various treatments, plants were transferred to 23 °C for 1 week. For the cold treatment, two-year-old pot-grown plants were subjected to a pre-chilled incubator at 4 °C, and the third and fourth fully expanded young leaves beneath the apex were sampled at the indicated time points. For the heat shock treatment, plants were placed at 42 °C in an incubator for the indicated time points. For salt stress treatments, plants were irrigated with 300 mM NaCl, and with sterile water for control and they were grown under normal conditions as the controls for the cold, heat and salt treatments. For the hormone treatments, leaves from the cutting plants were sprayed with 100 μM ABA, 0.5 g L⁻¹ methyl jasmonate (MeJA), or 100 μM salicylic acid (SA). Leaves sprayed with sterile water were used as a controls. Samples were collected at the indicated time points and immediately frozen in liquid nitrogen, prior to RNA extraction. Four plants were collected and pooled for each time point, and the sampling was performed in triplicate for biological replicates.

For organ-specific expression analysis, young/old leaves, tendrils, stems, and roots were collected from a 6-year-old *V. vinifera* cv. Cabernet Sauvignon grown under normal cultivation conditions at the Yuquanying experimental field at Ningxia University.

Real-time quantitative RT-PCR

Total RNA was extracted with the E.Z.N.A.® Plant RNA Kit and the extracts treated with DNase I (Omega) to remove genomic DNA. The RNA samples were converted to cDNAs

using the PrimeScript™ II 1st Strand cDNA Synthesis Kit (TaKaRa). qRT-PCR was carried out in 96-well reaction plates on a iQ5 PCR machine (Bio-Rad) in triplicate as previously described (Xu et al. 2014), with the grape GAPDH gene (GenBank accession No. EF192466) as the internal reference. For primer design, we extracted *MLP* sequences from assembled genome sequence of *V. vinifera* cv. Cabernet Sauvignon based on their scaffold location using a self-written perl script. The primers used for qRT-PCR analysis are listed in Table S1. Relative gene expression values were calculated using the $2^{-\Delta\Delta C_t}$ method (Livak and Schmittgen 2012). The generated data was graphically shown in visualized using a heatmap tool on OmicShare Tool (<http://www.omicshare.com/tools>).

Statistical analysis

The statistical analysis was carried out using the Statistical software SPSS Version 11.0. The data are presented as mean standard error of the mean (SE) and were tested for statistical significance by paired Student's test. $P < 0.05$ was selected as the point of minimal statistical significance in all analyses.

Results

Phylogenetic analysis of Bet v 1 allergen domain-containing proteins in plants

To identify and classify Bet v 1-related structures, we retrieved 1654 Bet v 1 allergen domain-containing protein sequences from the Pfam database. These were classified into five families, including MLP-like proteins, ABA receptors, S-noroclaurine synthases, major allergen Pru av 1/pathogenesis-related protein STH-2 s, and uncharacterized proteins in plants (Fig. 1). The most widely distributed group is MLP-like protein in plants. In particular, several members from grapevine cover all classified groups, with 14 members both in MLP-like proteins and the major allergen pru av 1 proteins, 3 members in each of the ABA receptor protein and pathogenesis-related protein STH-2 families, 2 members in the uncharacterized protein family, and 1 member in the S-noroclaurine synthase family (Fig. 1).

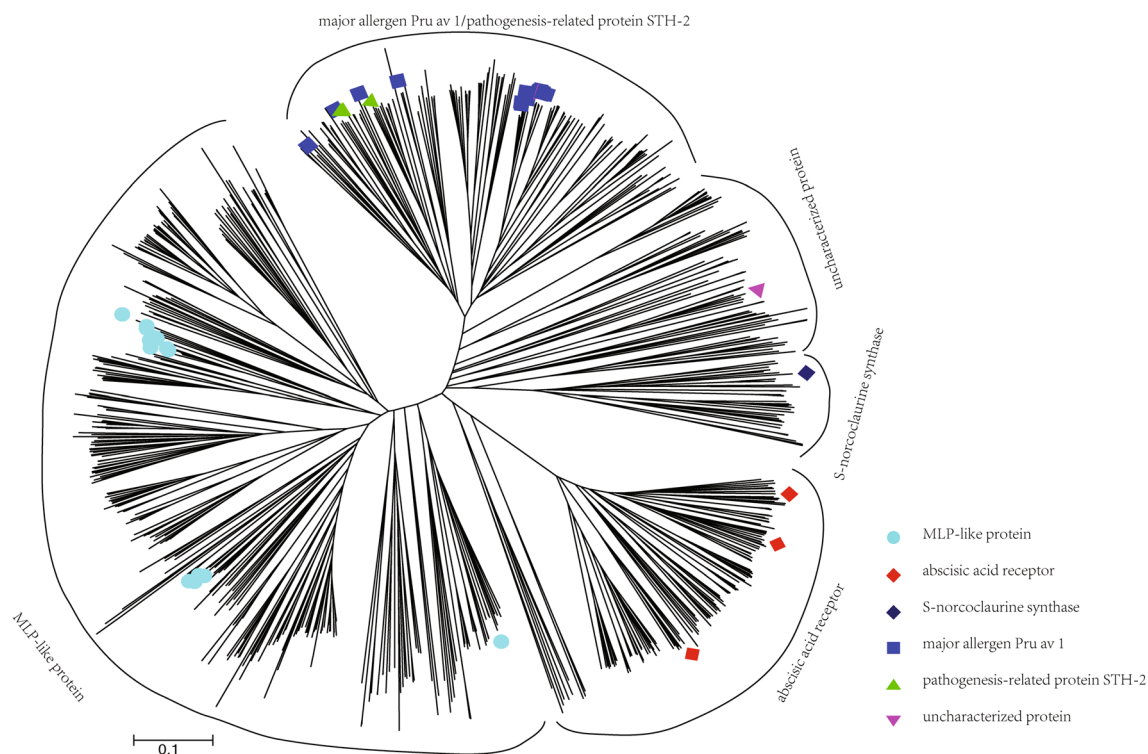


Fig. 1 Unrooted phylogenetic tree of 1 654 Bet v 1 allergen domain-containing proteins in plants. The amino acid sequences of the proteins were aligned with Clustal X version 2 using default settings. The phylogenetic tree was constructed using the NJ-method in the MEGA 6.0 software. The scale bar indicates the number of amino acid substitution per site. The five major lineages shown within line

enclosures are those of MLP-like proteins, abscisic acid receptors, S-noroclaurine synthases, major allergen Pru av 1/pathogenesis-related protein STH-2s, and uncharacterized proteins. The distribution of the 14 MLP-like proteins identified from grapevine, and their homologs from five major clusters, are indicated with different colored shapes

Distribution of MLP proteins in the plant kingdom

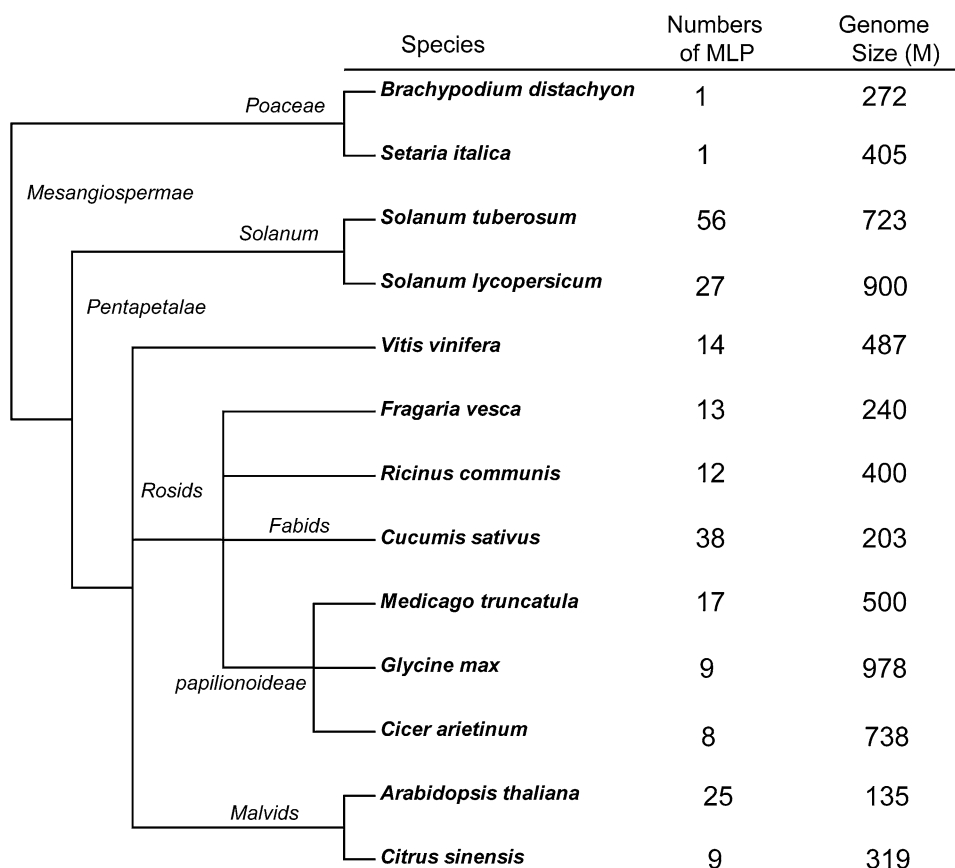
To expand the knowledge of MLP-like protein structures beyond those identified in grape species, 13 plant species whose genomes have been sequenced, including two monocots and 11 dicots were selected for further analyses. We identified *MLP* genes in each of the tested species, although the number varied substantially (Fig. 2). We observed that a species' genome size doesn't depend on the number of *MLP* genes it contains. In dicots, *Glycine max* possesses a large genome size (~978 M) but only 9 *MLP* genes exist in its genome, whereas *A. thaliana* with 135 M of genome size contains 25 *MLP* genes. Interestingly, only a few numbers of *MLP* homologs were found in monocots, while dicots had larger numbers (Fig. 2). For instance, it contains up to 56 *MLP* genes in *Solanum tuberosum*, and 27 *MLPs* in *Solanum lycopersicum*, whereas only 1 member of *MLPs* was found in monocots, *Brachypodium distachyon*, and *Setaria italica*. However, there is also widespread variation in copy number in *MLP* gene families in dicots plants (Fig. 2).

Chromosomal distribution, and phylogenetic analysis of duplication events

We further investigated the distribution of the 14 *VvMLP* genes in the grapevine genome, and found that they were all localized on chromosome 1, with 13 on the top arm and 1 on the bottom arm (Fig. 3a). We numbered the genes *VvMLP1* to *VvMLP14* according to their position on the chromosome, regardless of the strand on which they were located (Fig. 3a). Characteristics of the 14 *VvMLP* genes are shown in Table 1.

We next analyzed the phylogenetic relationship between grape and *A. thaliana* *MLP* sequences to gain insights into their origin and evolution. The 14 *Vitis* and 19 *A. thaliana* *MLP* genes clustered into three groups (Fig. 3b). The 6 genes in group I were exclusively from *V. vinifera*, whereas 14 genes in groups II were exclusive from *A. thaliana*. A total of 8 *VvMLP* genes and 11 *A. thaliana* *MLP* genes clustered within group III (Fig. 3b). The phylogenetic analysis suggests that there was duplication of *MLP* genes prior to the divergence of the last common ancestor of grape and *A. thaliana*. We also observed high levels of protein sequence similarity between the 14 *VvMLP* and 25 *A. thaliana* homologs. For instance, the *VvMLP14* and *AT1G24020* protein sequences share 56% amino acid similarity, whereas

Fig. 2 Summary of the evolution, numbers and genome size of MLP-like proteins in 13 plant species. The species tree of the MLP-like proteins from 13 plant species was generated via the online NCBI tool. The total numbers and genome size of *MLP* homologous genes found in each species is indicated



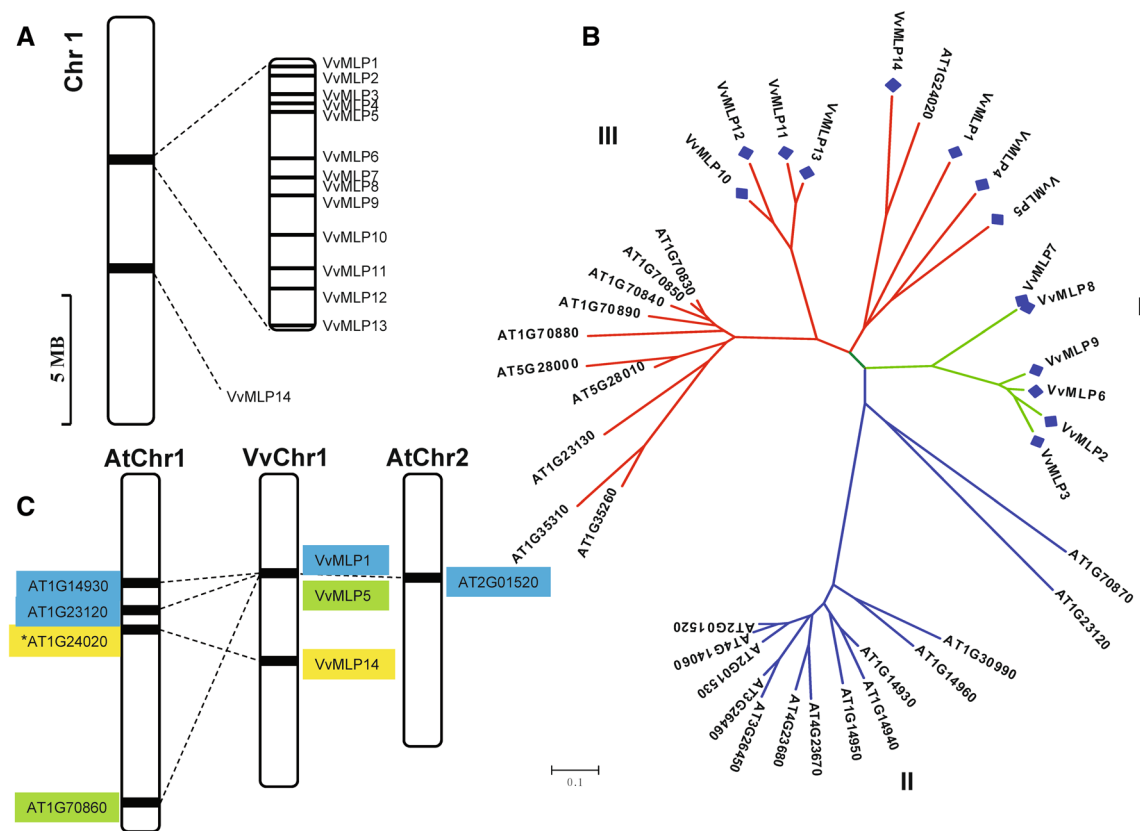


Fig. 3 Chromosomal location, phylogenetic relationship and synteny analysis of *VvMLP* genes from *V. vinifera* and homologs from the model plant species *A. thaliana*. **a** Chromosomal position of each *VvMLP* gene mapped onto the *V. vinifera* cv. Pinot Noir genome. The chromosome number is indicated. Nucleotide positions are referenced at Grape Genome Browser (12X). **b** Phylogenetic relationship of *MLP* proteins from *V. vinifera* and *A. thaliana*. The phylogenetic

trees were constructed based on NJ-algorithm with 1,000 bootstrap replicates. **c** Synteny analysis of *MLP* genes from *V. vinifera* and *A. thaliana*. Chromosomes containing *MLP* genes in either *V. vinifera* (*VvMLP*) or *A. thaliana* (*AtMLP*) are shown. The positions of the *AtMLP* and *VvMLP* genes are marked with black rectangles. Colored rectangles denote the syntentic relationships between *VvMLP* and *AtMLP* genes

Table 1 Characteristics of *MLP* family members in *V. vinifera*

Gene Name	Location	PN40024 12X V1 ID	ORF (aa)	pI/Mw(Kda)
<i>VvMLP1</i>	chr1 (+)4722643..4723231	VIT_01s0011g05060	150	5.27/16750.26
<i>VvMLP2</i>	chr1 (+)4725295..4726283	VIT_01s0011g05070	161	5.05/18212.88
<i>VvMLP3</i>	chr1(+)4729364..4737012		224	6.67/25334.24
<i>VvMLP4</i>	chr1 (+)4740067..4740955	VIT_01s0011g05090	151	6.17/16968.54
<i>VvMLP5</i>	chr1(+)4742652..4743585	VIT_01s0011g05100	151	5.32/16942.61
<i>VvMLP6</i>	chr1 (+)4757217..4758358	VIT_01s0011g05120	151	5.23/17117.13
<i>VvMLP7</i>	chr1(+)4765645..4767172	VIT_01s0011g05130	157	5.87/17953.29
<i>VvMLP8</i>	chr1 (+)4765555..4767224	VIT_01s0011g05130	151	5.34/17208.38
<i>VvMLP9</i>	chr1 (+)4774561..4775631	VIT_01s0011g05140	151	5.35/17080.14
<i>VvMLP10</i>	chr1 (+)4790431..4792324	VIT_01s0011g05150	151	5.62/17199.72
<i>VvMLP11</i>	chr1 (+)4810603..4813432	VIT_01s0011g05170	151	5.47/17027.75
<i>VvMLP12</i>	chr1 (+)4819645..4821386	VIT_01s0011g05180	151	5.30/17187.68
<i>VvMLP13</i>	chr1(+)4838740..4841982	VIT_01s0011g05190	151	5.63/17080.64
<i>VvMLP14</i>	chr1(-)9337035..9338314	VIT_01s0026g00570	153	5.16/17233.62

those of VvMLP10 and VvMLP13 share 52% amino acid similarity with AT1G70850 (Table S2).

An analysis of the segmental duplication showed no duplication events in the grape genome. However, synteny analysis between grape and *A. thaliana* demonstrated that the homolog of 3 grape *MLPs* (VvMLP1, VvMLP5, and VvMLP14) was present in the corresponding syntenic blocks of *A. thaliana*. As shown in Fig. 3c, a few putative segmental duplication events were shared between VvMLP5 and AT1G70850, VvMLP14 and AT1G24000, and VvMLP1 and AT1G14930/AT1G23120/AT2G01520. Interestingly, the VvMLP14/At1G24000 pair showed a syntenic relationship but At1G24000 was not observed in our phylogenetic tree (Fig. 3b). When we looked further into this locus, we found that At1G24000 also contains the Bet v 1 domain and shares a 33% similarity with VvMLP14 at the amino acid level (Table S2). However, our phylogenetic analysis indicated that the phylogenetically close VvMLP14/At1G2420 pair, the proteins with which At1G24000 shares a 56% amino acid similarity (Table S2), have a closer syntenic relationship. Based on the phylogenetic analysis and syntenic relationship

analysis, we chose to annotate VvMLP14 as being syntenic with At1G2420.

Characteristics of VvMLP domains

To further explore the phylogenetic relationship between the 14 grape MLP proteins, a neighbor-joining (NJ) phylogenetic tree was constructed. The phylogenetic tree showed that the protein sequences could be categorized into three major groups: (1) VvMLP2/3/6/7/8/9, (2) VvMLP10/11/12/13, and (3) VvMLP1/4/5/14 (Fig. 4a). Predicted proteins in group I show a high degree of similarity (60–85%), members in group II shared a moderate degree of similarity (68–78%), while members of group III have a lower degree of similarity (36–47%). It is worthwhile mentioning that two adjacent gene members VvMLP7 and 8 shared 100% amino acid identity (Table S3).

Analysis of the exon/intron structures indicated that most of the coding sequences were disrupted by introns (Fig. 4b). 12 of the *MLP* genes have two exons, whereas two VvMLP genes (2 and 3) harbor 3 exons. Most genes that clustered

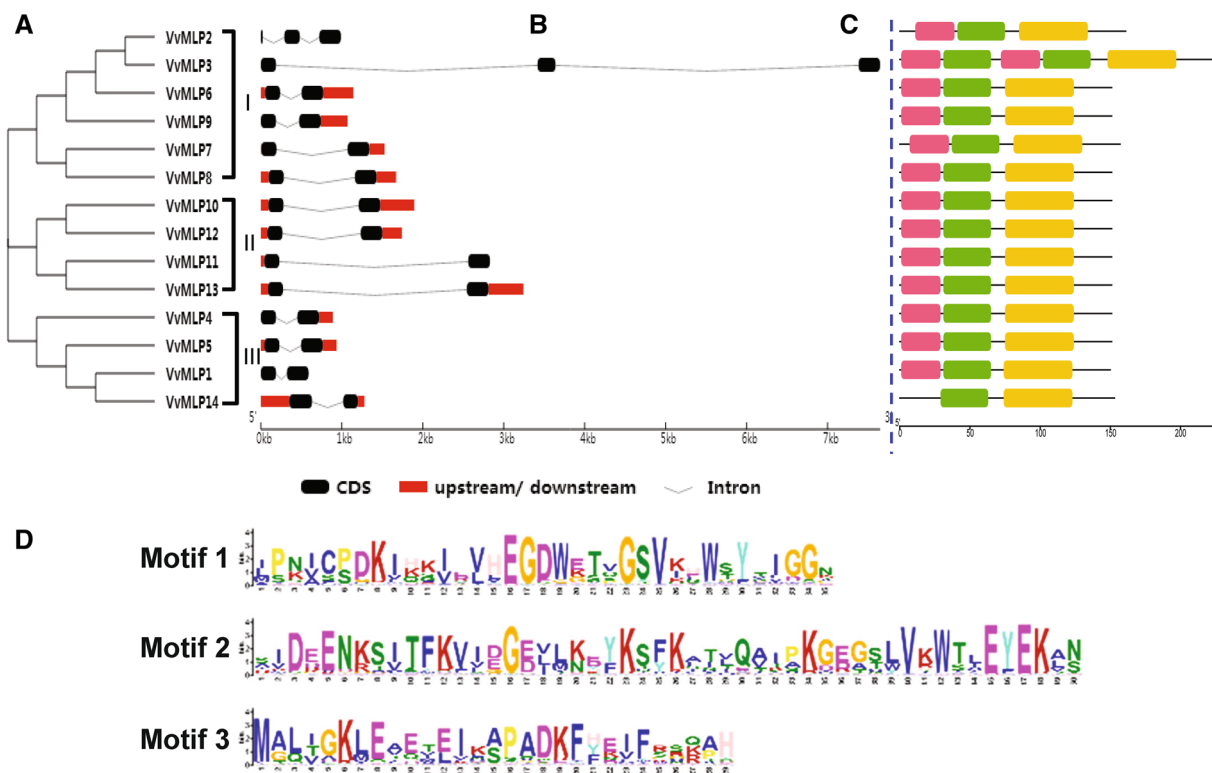


Fig. 4 Phylogenetic tree, gene structure, and architecture of conserved motifs in deduced VvMLP proteins. **a** The tree includes 14 MLP proteins from *Vitis vinifera*. The phylogenetic tree was constructed using MEGA 6 and the NJ-algorithm with 1000 bootstrap replicates. Subgroups within line enclosures are indicated. **b** Exon/intron organization of VvMLP genes. Exon(s), intron(s) and upstream/downstream regions are represented by black rectangles, grey lines

and red rectangles, respectively. **c** Architecture of conserved protein motifs in the VvMLP family. **d** Sequence logos of 3 conserved motifs detected in the VvMLP proteins. The overall height of each stack represents the degree of conservation at this position, while the height of the individual letters within each stack indicates the relative frequency of the corresponding amino acid

in the same subgroup exhibited similar exon/intron structures, but the length of the introns was distinct. In particular, *VvMLP3* possessed two long introns (3.2 and 3.7 kb, respectively), which differed from the other genes.

To identify structural features in the 14 predicted *VvMLP* proteins, MEME was used for motif searching, and sequence logos of motif 1, 2, and 3 were drawn to determine the conserved amino acid residues (Fig. 4c). Combined block diagrams of signatures generated on the basis of the MEME analysis showed that protein architectures are broadly conserved in the *VvMLP* family, and three motifs are prevalent in 12 out of 14 *VvMLP* proteins (Fig. 4d). As an exception, motif 3 is absent from the N-terminal region of *VvMLP14*. Additionally, *VvMLP3* has two repeats of motif 1 and 3 towards its N-terminal, rather than having all the 3 motifs in a pattern similar to that in the other 12 *VvMLP* proteins.

Comparative analysis of *MLP* members between Pinot Noir and Cabernet Sauvignon

With the information of the *V. vinifera* cv. Pinot Noir genome (Jaillon et al. 2007) and the assembled sequence of *V. vinifera* cv. Cabernet Sauvignon genome (Chin et al. 2016), we extracted and compared the *MLP* members and their promoter sequence in two *Vitis* species. Our results showed that these two European *Vitis* species have very high identities for the corresponding CDS (Table S4) and promoter regions (Table S5). Interestingly, we noted that several pairs of *VvMLP* genes from Cabernet Sauvignon share the same promoter sequences, such as *VvMLP2/3/6*, *VvMLP4/5*, *VvMLP7/8*, and *VvMLP11/12/13*. Similarly,

MLP members from Pinot Noir display the same patterns, and share highly similarities in promoter sequences for each variety (Table S5). This finding allowed us to determine the roles of these gene members in response to environmental stress in *V. vinifera* cv. Cabernet Sauvignon. Several members of the *VvMLP* family showed a high nucleotide sequence identity between each other, which limits the design of gene-specific primers for each member. We therefore, designed seven pairs of primers (Table S1) to examine the expression of seven groups of *VvMLP* genes, either in different tissue/organs or in response to environmental stimuli (*VvMLP1*, *VvMLP2/3/6/9*, *VvMLP4*, *VvMLP5*, *VvMLP7/8*, *VvMLP11/12/13*, and *VvMLP14*).

In silico analysis of *cis*-acting regulatory elements in the putative promoter regions in Cabernet Sauvignon

To identify potential *cis*-acting regulatory DNA sequences responsible for abiotic or biotic stress responses in the putative promoter region of the *VvMLP* genes, we used 1.0 kb upstream sequences from the 14 *VvMLP* genes, extracted from assembled genome sequence of *V. vinifera* cv. Cabernet Sauvignon, and performed a search of *cis*-regulatory elements with MatInspector program. Due to the quality of the assembled data of Cabernet Sauvignon, only 417-bp 5'UTR regions of *VvMLP11/12/13* genes could be extracted. The in silico analysis of *VvMLP* upstream regions from Cabernet Sauvignon led to the identification of several putative *cis*-acting regulatory elements associated with abiotic and biotic stress responses (Fig. 5). The abiotic stress-related elements

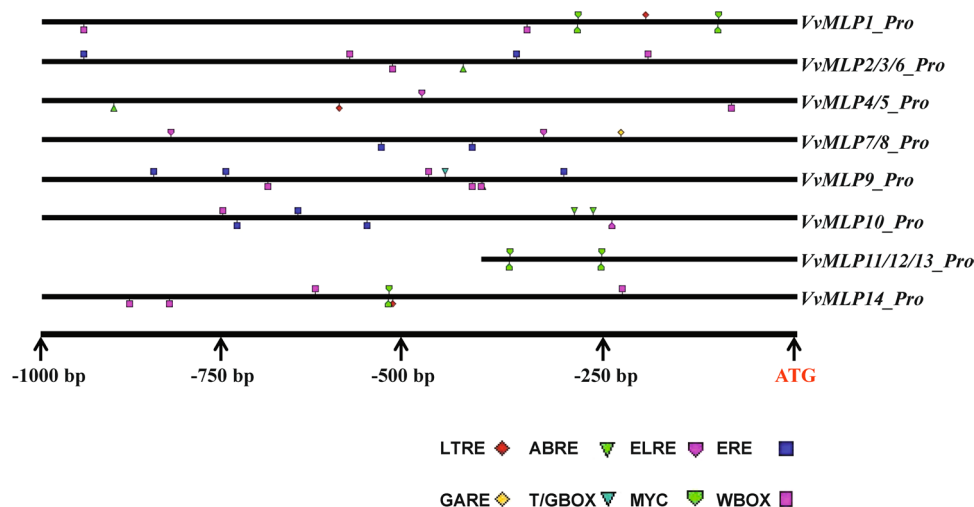


Fig. 5 Putative *cis*-acting elements present in *VvMLP* gene promoters. Positions of the different *cis*-acting elements within a 1 kb region upstream from the translation start codon (ATG) are indicated by color-coded shapes. Vertical arrows indicate the positions upstream of the ATG. LTRE, low temperature responsive element; ABRE, a

cis-acting element involved in response to abscisic acid; ELRE, elicitor-responsive element; EREL, ethylene responsive element; GARE, gibberellin-responsive element; T/GBOX, a *cis*-acting element involved in jasmonate (JA) induction; MYC, dehydration-responsive element; WBOX, fungal elicitor-responsive element

included low-temperature responsive element (LTRE), and dehydration-responsive elements (MYC). In addition, elements involved in biotic stress were elicitor-responsive elements (WBOX, and EREL). Some phytohormone regulatory elements were also identified in several *VvMLP* upstream regions, such as abscisic acid responsive elements (ABRE), ethylene responsive element (ERE), gibberellin-responsive element (GARE), and jasmonate (JA) responsive element (T/G-box).

Organ-specific expression pattern of *VvMLP* genes

To explore the organ-specific presence of transcripts from the 14 *VvMLP* genes in more detail, we performed a global

analysis of the grapevine (cv. Corvina) transcriptome with 54 different organs/tissues at various developmental stages (Table S6) using microarray data retrieved from Grape eFP Browser microarray database. The result from hierarchical clustering analysis demonstrated that members clustering in the same group shared a similar expression pattern in various organs/tissues during development (Fig. 6a). It is worth to note that *VvMLP2*, showed relatively high and constitutive expression in all organs/tissues during development. A large group, containing ten members (*VvMLP1/4/5/6/7/8/9/10/12/13*) displayed large variations in expression profiles, of which *VvMLP4* exhibited high levels of expression in leaves (fruit set and senescent), whereas *VvMLP12/13* showed high levels of expression in

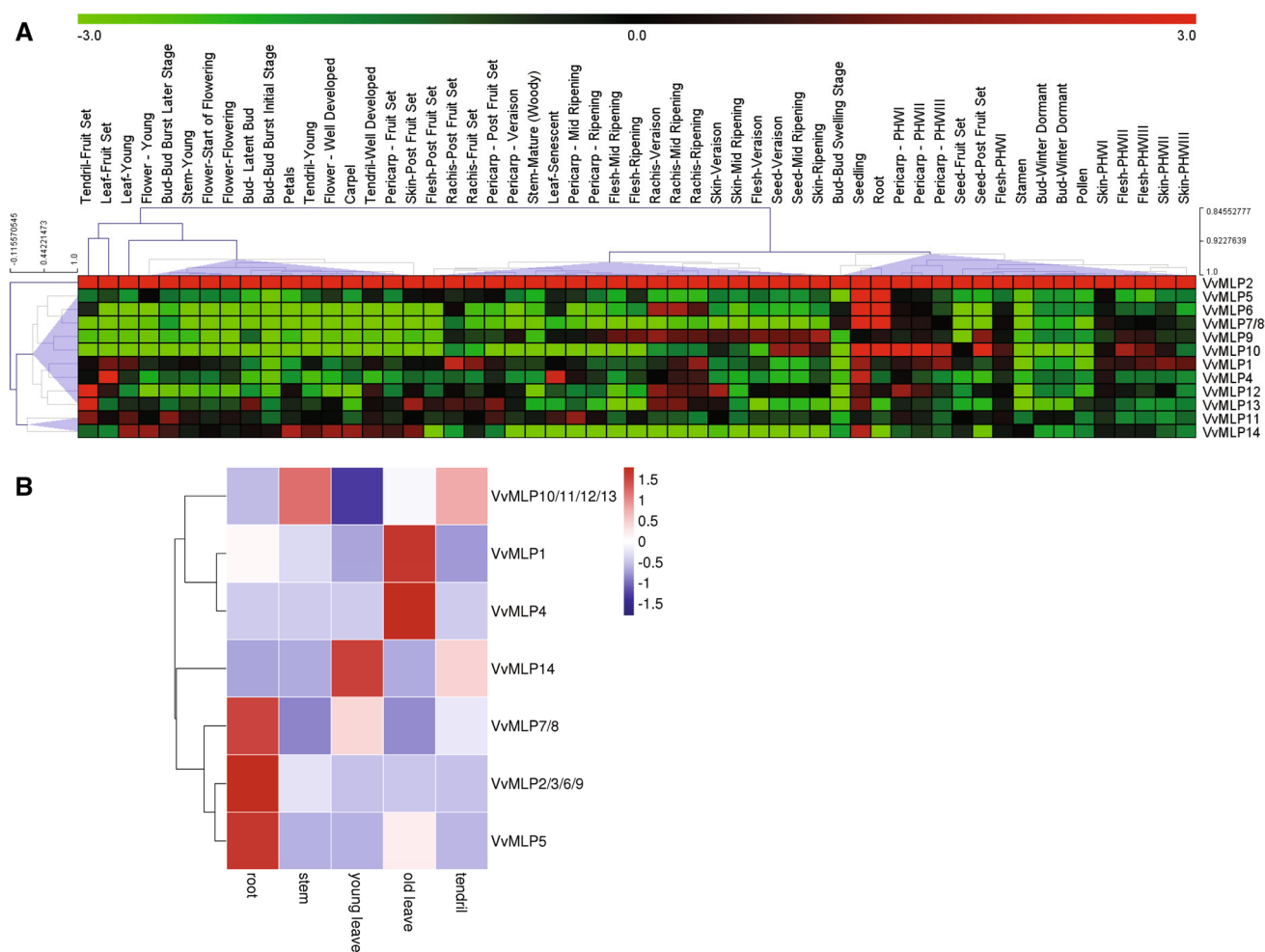


Fig. 6 Expression profiles of the *VvMLP* genes in different grape organs, tissues and developmental stages. **a** Expression of *VvMLP* genes in the *V. vinifera* cv ‘Corvina’ atlas (http://bar.utoronto.ca/efp_grape/cgi-bin/efpWeb.cgi). Data were Z-core normalized based on the mean expression value of each gene in all tissues analyzed. Genes were hierarchically clustered based on the average Pearson’s distance metric and the ‘average linkage’ method. Red and green boxes indicate high and low expression levels, respectively, for each

gene. **b** Expression profiles of *VvMLP* genes in roots, stems, young leaves, old leaves and tendrils from 6-year old *V. vinifera* cv. Cabernet Sauvignon cultivated under field conditions. Heatmap representation of the qRT-PCR data for the 14 *VvMLP*s in five different tissues was produced and graphically visualized using the OmicShare platform. Orthologs with very high identities that could not be distinguished by gene-specific primers were examined together use the same primers

tendrils (fruit set). Other members of this group, such as *VvMLP5/6/7/8*, exhibited high levels of expression in seedlings and roots. Importantly, *VvMLP10* displayed a relatively high level of expression in many tissues/organs (seedling, root, pericarp (WI and WII), and seed (fruit set and post-fruit set)).

We next designed primers according to the extracted sequence from the assemble Cabernet Sauvignon to analyze the expression pattern of *VvMLPs* in young leaves, old leaves, tendrils, stems and roots from *V. vinifera* cv. Cabernet Sauvignon (Fig. 6b). The high similarity in nucleotide sequences made it difficult to design primers to distinguish all the *VvMLP* genes. We therefore, used seven pairs of primers to cover the 14 *VvMLP* genes, with *VvMLP2/3/6/9*, *VvMLP7/8*, or *VvMLP10/11/12/13* sharing the same primer sets. The expression analysis revealed that the genes were expressed highly in specific organs (Fig. 6b), partially similar to the atlas of several members in Corvina (Fig. 6b). For instance, *VvMLP2* displayed a high transcript abundance in old leaves, but a low expression levels in other tested tissues in Cabernet Sauvignon, which is different from Corvina (Fig. 6a). However, *VvMLP 4* in Cabernet Sauvignon showed a high transcript abundance in old leaves, similar to those observed in Corvina.

Expression profiles of *VvMLP* genes in response to abiotic stresses and phytohormone treatments

The *cis*-acting regulatory element analysis revealed that the promoters of the *VvMLP* genes harbor several putative *cis*-acting regulatory elements involved in abiotic/biotic stress and phytohormone responses. We therefore, speculated that their expression is induced in response to multiple environmental stimuli. Using qRT-PCR, we assayed the transcript abundances of 14 *VvMLP* genes in Cabernet Sauvignon following exposure to three types of abiotic stress (cold, heat, and high NaCl concentrations), as well as three phytohormone treatments (ABA, MeJA and SA) at 12 and 24 h. The expression profiles of the *VvMLP* family genes under various stresses and treatments are shown in Fig. 7.

Under cold stress, *VvMLP1*, *2/3/6/9*, *4*, and *7/8* rapidly increased at 12 h but decreased at 24 h (Fig. 7). We observed no changes in the expression of *VvMLP10/11/12/13* upon cold stress, either at 12 or 24 h, whereas a relatively high expression level was detected in *VvMLP5* after 48 h of cold stress. More than half of the *VvMLP* genes were either down-regulated or showed no change in expression under more than one stress conditions. However, many were up-regulated by heat stress. For example, *VvMLP2/3/6/9* and *VvMLP7/8* showed a high up-regulation under heat stress either at 12 or 24 h. In addition to salt stress, *VvMLP1*, *2/3/6/9*, *4*, and *7/8* were up-regulated upon 48 h of NaCl

stress, whereas *VvMLP5*, and *14* were down-regulated by salt stress (Fig. 7).

Similarly, the expression patterns of the 14 *VvMLP* genes in response to phytohormone treatments varied. As shown in Fig. 7, 6 *VvMLP* genes (*VvMLP5*, *10/11/12/13*, and *14*) were repressed by ABA after 12 h of treatment, 4 members (*VvMLP1*, *4*, and *7/8*) exhibited increased expression levels, whereas the other genes (*VvMLP2/3/6/9*) showed no obvious change in expression (Fig. 7). In contrast, the expression of 4 of the 14 *VvMLP* genes (*VvMLP1*, *4*, and *7/8*) was up-regulated within 12 h treatment with MeJA. Only 4 *VvMLP* genes exhibited increased expression levels in response SA treatments, while the others showed either down-regulated profiles or no obvious changes. These results suggest that the *VvMLP* genes respond to environmental stimuli, which may be linked to stress-induced alterations in phytohormone levels.

Discussion

In this study, we presented genome-wide evolutionary characterization and expression analysis of *MLP* genes in *V. vinifera* in response to various abiotic stresses. We determined that there are 14 *MLP* genes in the *V. vinifera* genome. We also showed that the promoter-sharing by several *MLP* genes with the same phylogenetically cluster was present in the chromosome to a greater extent than expected by chance. We confirmed the superfamily organization, evolution and *cis*-elements in the upstream region. Their tissue-specific expression and involvement in responses to environmental, and hormonal stress are further supported by current results. Our data also provides a foundation for further functional investigations of these *MLP* genes and potential to be used in abiotic/biotic tolerant improvements.

Evolution of plant *MLP* proteins

There is an increasing amount of evidence supports a linear relationship between the size of genome and number of genes in prokaryotes (Lynch and Conery 2003; Konstantinidis and Tiedje 2004). However, when it comes to eukaryotes (e.g., plants), this relationship does not seem to exist. This is also supported by the fact that the numbers of *MLP* protein largely varied in plants regardless of the difference in size of their genomes (Fig. 2). Two possibilities exist. In eukaryotes, most of the DNA do not encode anything, including, for example, noncoding DNA sequences (Niu and Jiang 2013). Besides, a chromosome is only an individual molecule, and the number of chromosomes in an organism could not provide any information on the size of the DNA or the number of genes it contains.

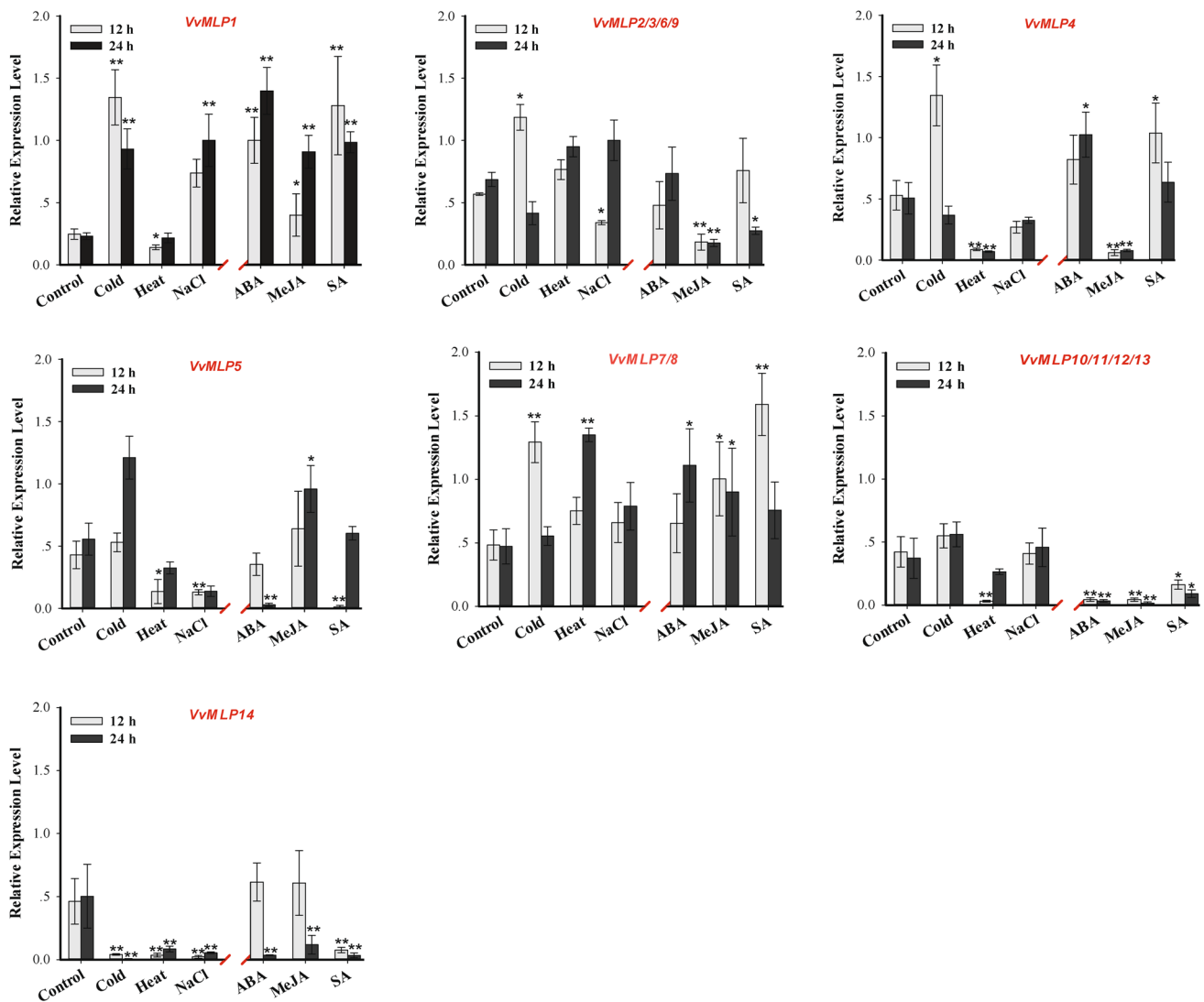


Fig. 7 Expression of *VvMLP* genes 12 h (gray bars) and 24 h (black bars) post-treatment in 2-year-old potted *Vitis vinifera* cv. Cabernet Sauvignon plants grown under control conditions, or exposed to cold (4 °C), heat (42 °C), NaCl (300 mM), abscisic acid (ABA, 100 μM), methyl jasmonate (MeJA, 0.5 g/L) and salicylic acid (SA, 100 μM).

The expression of the *MLP* genes was measured by quantitative real-time RT-PCR. The data are presented as mean $\pm$ SE from three independent experiments. * P < 0.05, ** P < 0.01 versus values of *VvMLP* expression in *V. vinifera* cv. Cabernet Sauvignon in the control conditions either 12 or 24 h post-treatment

One of the most interesting findings in our analysis of *MLP* proteins in monocots and dicots was that the relatively small gene numbers encountered in monocot plants (Fig. 2). This finding was also documented in evolutionary analyses of ‘non-family’ (NF) genes in 14 diverse plants, where genes with only 1 or 2 copies per genome were termed as monocot-specific manner (Ye et al. 2013). The number of genes in monocot-specific manner was significantly lower than that in dicot plants. One potential explanation is that NF gene resulted largely from a higher rate of gene loss after segmental duplications and/or lower frequency of tandem duplications (Ye et al. 2013). Consistent with this idea, several lines of evidences indicated that tandem duplication

contributes to the expansion of gene families (Cannon et al. 2004; Yang et al. 2008), and gene loss is a common aspect of evolutionary dynamics after gene duplication (Yang and Tuskan 2006; Freeling 2009).

Functional categories of *MLP* genes with other *Bet v* 1-containing protein in *Vitis* plants

It has been reported that the *Bet v* 1 fold is an ancient, versatile scaffold for binding of large, hydrophobic ligands (Radauer et al. 2008). Also, previous reports indicated that the plant-specific *MLPs* belong to the *Bet v* 1 protein family (Yang et al. 2015). To date, there are no reports on the

biological roles of *MLP* members in grapevine. The released complete genome of *V. vinifera* cv. Pinot Noir (PN40024) (Jaillon et al. 2007) and assembled sequence from *V. vinifera* cv. Cabernet Sauvignon (Chin et al. 2016), makes it possible to allow us to characterize the *MLP* gene family and compare their gene structures, evolution, duplication event and expression profiles in response to various environmental stresses. Our phylogenetical analysis within plant Bet v 1-containing proteins, yielded a classification into five groups (Fig. 1). We noted that several members of the Bet v 1 related structure were found in *Vitis* genome. Interestingly, although the predicted *MLP*-like proteins contain the allergen Betv 1 domain, they are distinct from proteins in the other clusters (Fig. 1). It is generally accepted that genes belonging to the same functional group tend to have similar expression patterns and share expression regulation mechanisms (Siu-Ming et al. 2008; Yang et al. 2015).

Several other proteins were clustered into the group of major allergen Pru av 1/pathogenesis-related protein STH-2s (Fig. 1), among which PR10 proteins are represented members involved in various pathogens, low temperature, UV radiation, or oxidative stress (Hammondkosack and Jones 1996; Lcvan and Eavan 1999; Marković-Housley et al. 2003; Van et al. 2006). One Bet v 1-related protein in grape was found in group of S-noroclaurine synthase, which is the only known member of the PR10/Bet v1 family shown unequivocally to possess catalytic activity in plants (Lee and Facchini 2010). ABA receptors are considered the major regulators of ABA-dependent gene expression (Guo et al. 2011b), participating in various abiotic stress response in rice (Ye et al. 2012; He et al. 2014). The occurrence of three members from grape indicated that the ABA receptors are highly conserved in dicot and monocot plants evolutionarily. Similar to the structure of an uncharacterized protein from *Arabidopsis* (Radauer et al. 2008), one member from grape also falls into this group (Fig. 1). These results suggest the diverse roles of Bet v 1-containing proteins in grapevine.

Co-expression of genes and its functional implications

A further focus on the 14 identified *MLPs* from *V. vinifera* revealed that they could be classified into three clusters (Fig. 4a). We therefore, infer that these *MLPs* in the same cluster have similar functions. Interestingly, we found the genomic arrangement of the 14 *MLPs* in *Vitis* genome is relatively uncommon, which are predominantly arranged in tandem clusters of chromosome 1 (Fig. 3a). What is surprising is that members of the *V. vinifera* pathogenesis-related 10 (PR10)-like protein family are also located in a single cluster on chromosome 5 (Lebel et al. 2010). Whether the other Betv 1-related proteins with same distribution feature are still required to be determined.

Importantly, we note that promoter-sharing by several members with high level of coding region sequence identity in the same cluster, e.g., *VvMLP2/3/6*, *VvMLP4/5*, *VvMLP7/8* and *VvMLP11/12/13* (Table S3 and S5). Particularly, two adjacent members of *VvMLP7/8* are identity in both the coding regions and promoter region sequences but the arrangement location in chromosome is different. For gene pair of *VvMLP2/6*, they share relatively high identity in the coding regions except for *VvMLP2* has a 30-bp longer in 5' exon 1 than that of *VvMLP6*, suggesting alternative splicing event occurred between these two gene. The promoter-sharing by different genes and the conservation of this phenomenon are reported on *CPNE1* and *RBM12* of human genome (Siu-Ming et al. 2008). It is possible that this arrangement reflects a selection constraint to keep these genes co-regulated during evolutionary courses (Siu-Ming et al. 2008).

Expression patterns of *VvMLPs* in different tissues

The diverse expression patterns of genes in different plant tissues may provide comprehensive insights into the genetic basis of plant. We, therefore, investigated expression profiles of *MLP* genes in various organs of two cultivars from both Corvina and Cabernet Sauvignon. A comprehensive analysis of the global grape (cv. Corvina) transcriptome profiles of various tissues/organs at different developmental stages was previously reported (Fasoli et al. 2012). We used this information, focusing on the genes corresponding to the 14 *VvMLP* genes, and analyzed their transcript profiles using the Grape eFP Browser microarray database. We found the *VvMLP* genes to be expressed in different tissues in *V. vinifera* cv. Corvina, suggesting that the physiological roles of these gene may be distinct. We observed that most of the *VvMLP* genes in Cabernet Sauvignon showed similar patterns to those available in the Corvina microarray databases (Fig. 6a, b).

However, we observed that not all members between two cultivars have the same organ-specific expression pattern. *VvMLP2* expression was constitutively high in all tissue/organs at different developmental stages in Corvina (Fig. 6a) but was preferentially expressed in the roots of Cabernet Sauvignon (Fig. 6b). Similar observation was reported on *C. pepo* ssp. *Pepo* where the expression levels of *MLPGR3* in Black Beauty, Gold Rush, Raven, and Magda, were much higher than in the Patty Green (Inui et al. 2013). It is generally accepted that gene expression is influenced by multiple factors, such as polymorphisms, the genomic backgrounds and environment. A recent *De novo* transcriptome revealed the characterization of varietal diversity between *V. vinifera* cv. Corvina and the PN40024 reference genome (Venturini et al. 2013). Thus, this discrepancy could be explained, in part, by assuming that some of the *MLP* genes

were expressed in a variety-dependent manner. Although a detailed investigation of these factors is beyond the scope of this work, we cannot rule out the possibility that primers designed on the basis of the *V. vinifera* cv. Pinot noir clone PN40024 genome sequence did not anneal to cDNA molecules derived from *V. vinifera* cv. Cabernet Sauvignon because of sequence divergence of these loci between the two grape genotypes. However, tissue-specific variations in transcript abundance among the *VvMLP* genes is consistent with previous results shown for the *MLP* gene family of *A. thaliana* (Chen and Dai 2010). Considering the data from both grape varieties, we concluded that different genes are expressed in specific tissues/organs, likely associated with gene functional differentiation.

Functions of *VvMLPs* in response to various abiotic and hormone stresses

It has been reported that a peach *MLP* gene is involved in fruit and flower development and in pathogen defense responses (Ruperti et al. 2002). In *A. thaliana*, the expression of *MLP* genes was shown to be induced by a variety of environmental stimuli or stresses (Chen and Dai 2010), and the expression of two members with high sequence similarity, *AtMLP28* (At1G70830) and *AtMLP43* (At1G70890), as well as *At4G23670* was induced by gravity or mechanical stimulus (Kimbrough et al. 2004). *At4G23670* expression can be observed only during germination and seedling development, and only in flowers (Guo et al. 2011a). In contrast, *At4G23780* shows continuous expression throughout development in all tissues. Although the biological roles of the *MLPs* is still unknown, these investigations suggest that the roles of the expression patterns were involved in defense or stress responses.

To gain a better understanding of how these *VvMLPs* respond to environmental stresses, the *cis*-regulatory elements analysis was determined over the upstream regions of the gene family in Cabernet Sauvignon. Our analysis revealed that there are several *cis*-acting regulatory elements associated with biotic/abiotic stress and phytohormone responses (Fig. 5). LTRE (Low temperature Responsive Element), a *cis*-element was reported in response to low temperature in *Arabidopsis* (Baker et al. 1994). We found this *cis*-element was present in the promoter regions of four *VvMLP* members (*VvMLP1*, 4/5, and 14), and the cold-inducible profile occurred with either up-regulation or down-regulation in the four genes at either 12 or 24 h (Fig. 7).

Interestingly, we observed that several members (e.g., *VvMLP2/3/6/9*, and 7/8 at 12 h) were involved in cold-induced expression, although there were no such LTRE elements detected in their upstream regions (Fig. 6). Similarly, ABRE, an ABA-responsive element (Kuroi and

Nakamoto 2003), was detected in the upstream region of *VvMLP2/3/6*, 4/5, 9, and 10 but not in the others. However, these members were still responded to ABA (Fig. 7). This finding was unexpected and suggests that the presence or absence of *cis*-acting elements in the promoter region is not directly related to the corresponding stress expression pattern. Contrary to *VvMLP14*, *VvMLP1* is transcriptionally up-regulated by various tested stresses, which suggests the two genes are simultaneously activated under an abiotic stress condition could exert reverse roles in response to abiotic stresses. This finding also provides two candidate gene resources in which could be used for grape stress-resistant improvement through forward or reverse genetics, respectively. We also found most of the highly induced *VvMLP* genes have overlapping responses to salt, heat and cold stress, similar to observations from transcriptome analyses of *A. thaliana* responses to salt, osmotic, and cold stress (Kreps et al. 2002). The *VvMLP* genes also showed different expression profiles in response to phytohormone treatments, but all members responded to at least one phytohormone treatment, with the exception of *VvMLP11/12/13*. It is possible that these genes are independent of the phytohormone regulatory signaling pathway.

It remains unresolved, however, whether promoter-sharing by non-identical but highly similar members of *VvMLP* could play same biological roles. Our ability to better understand their potential roles would depend on detailed investigation of the functional roles and expression profiles of these genes in transgenic plants. Our data provide evidence that the *VvMLP* family is likely to be involved in responses to abiotic stresses, in addition to phytohormones, and their different expression suggests distinct roles during these responses.

In conclusion, to the best of our knowledge, this is the first comprehensive genome-wide analysis of the *MLP* gene family in *V. vinifera*. The *V. vinifera* genome has a family of 14 *VvMLP* genes, which are tandemly arranged and tightly paired on chromosome 1. Importantly, promoter-sharing by non-identical but highly similar members was found over these *VvMLP* family clustered single chromosome. Our findings may represent a new phenomenon in gene expression regulation. Several of them are transcriptionally active in different grapevine tissues, and some *VvMLP* genes such as *VvMLP1*, and 14 are good candidate genes for grape stress-resistant improvement. Our study provides a comprehensive understanding and facilitates further functional characterization of the grape *MLP* gene family in gene regulation networks of stress resistance and development.

Acknowledgements This study was funded by the Western First-class Discipline Construction Project of Horticulture (Grant Number

NYXLXK2017B03), the Major Science and Technology Program of Ningxia Hui Autonomous Region (Grant Number 2016BZ06) and National Natural Science Foundation of China (Grant Number 31560550).

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

Ethical standards This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Baker SS, Wilhelm KS, Thomashow MF (1994) The 5'-region of *Arabidopsis thaliana* cor15a has cis-acting elements that confer cold-, drought- and ABA-regulated gene expression. *Plant Mol Biol* 24(5):701–713
- Biais B, Krisa S, Cluzet S, Da Costa G, Waffo-Teguo P, Mérillon J-M, Richard T (2017) Antioxidant and cytoprotective activities of grapevine stilbenes. *J Agric Food Chem* 65(24):4952–4960
- Cannon SB, Mitra A, Baumgarten A, Young ND, May G (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol* 4(1):10
- Cheadle C, Vawter MP, Freed WJ, Becker KG (2003) Analysis of microarray data using Z score transformation. *J Mol Diagn* 5(2):73–81
- Chen J-Y, Dai X-F (2010) Cloning and characterization of the *Gossypium hirsutum* major latex protein gene and functional analysis in *Arabidopsis thaliana*. *Planta* 231(4):861–873
- Chin CS, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A, Dunn C, O'Malley R, Figueroabalderas R, Moralescruz A (2016) Phased diploid genome assembly with single-molecule real-time sequencing. *Nat Methods* 13(12):1050–1054
- Chruszcz M, Ciardiello MA, Osinski T, Majorek KA, Giangrieco I, Font J, Breiteneder H, Thalassinou K, Minor W (2013) Structural and bioinformatic analysis of the kiwifruit allergen Act d 11, a member of the family of ripening-related proteins. *Mol Immunol* 56(4):794–803
- Facchini PJ, Park S-U (2003) Developmental and inducible accumulation of gene transcripts involved in alkaloid biosynthesis in opium poppy. *Phytochemistry* 64(1):177–186
- Fasoli M, Dal Santo S, Zenoni S, Tornielli GB, Farina L, Zamboni A, Porceddu A, Venturini L, Bicego M, Murino V (2012) The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. *Plant Cell* 24(9):3489–3505
- Freeling M (2009) Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. *Annu Rev Plant Biol* 60:433–453
- Griffing LR, Nessler CL (1989) Immunolocalization of the major latex proteins in developing laticifers of opium poppy (*Papaver somniferum*). *J Plant Physiol* 134(3):357–363
- Guo D, Wong WS, Xu WZ, Sun FF, Qing DJ, Li N (2011a) Cis-cinnamic acid-enhanced 1 gene plays a role in regulation of *Arabidopsis* bolting. *Plant Mol Biol* 75(4–5):481–495
- Guo J, Yang X, Weston DJ, Chen JG (2011b) Abscissic acid receptors: past, present and future. *J Integr Plant Biol* 53(6):469–479
- Hammond-Kosack KE, Jones JD (1996) Resistance gene-dependent plant defense responses. *Plant Cell* 8(10):1773–1791
- He Y, Hao Q, Li W, Yan C, Yan N, Yin P (2014) Identification and characterization of ABA receptors in *Oryza sativa*. *PLoS One* 9(4):e95246
- Inui H, Sawada M, Goto J, Yamazaki K, Kodama N, Tsuruta H, Eun H (2013) A major latex-like protein is a key factor in crop contamination by persistent organic pollutants. *Plant Physiol* 161(4):2128–2135
- Jaillon O, Aury J-M, Noel B, Policriti A, Clepet C, Casagrande A, Choise N, Aubourg S, Vitulo N, Jubin C (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449(7161):463–467
- Kambiranda D, Katam R, Basha SM, Siebert S (2013) iTRAQ-based quantitative proteomics of developing and ripening muscadine grape berry. *J Proteome Res* 13(2):555–569
- Kimbrough JM, Salinas-Mondragon R, Boss WF, Brown CS, Sederoff HW (2004) The fast and transient transcriptional network of gravity and mechanical stimulation in the *Arabidopsis* root apex. *Plant Physiol* 136(1):2790–2805
- Konstantinidis KT, Tiedje JM (2004) Trends between gene content and genome size in prokaryotic species with larger genomes. *Proc Natl Acad Sci USA* 101(9):3160–3165
- Kreps J, Wu Y, Chang H, Zhu T, Wang X, Harper JF (2002) Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress. *Plant Physiol* 130(4):2129–2141
- Kuroi T, Nakamoto K (2003) Two different novel cis-acting elements of *erd1*, a *clpA* homologous *Arabidopsis* gene function in induction by dehydration stress and dark-induced senescence. *Plant J* 33(2):259–270
- Kutchan TM (2005) A role for intra- and intercellular translocation in natural product biosynthesis. *Curr Opin Plant Biol* 8(3):292–300
- Lcvan L, Eavan S (1999) The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiol Mol Plant Pathol* 55(2):85–97
- Lebel S, Schellenbaum P, Walter B, Maillot P (2010) Characterisation of the *Vitis vinifera* PR10 multigene family. *BMC Plant Biol* 10(1):184–184
- Lee EJ, Facchini P (2010) Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family. *Plant Cell* 22(10):3489
- Lee O-R, Sathiyaraj G, Kim Y-J, In J-G, Kwon W-S, Kim J-H, Yang D-C (2011) Defense genes induced by pathogens and abiotic stresses in *Panax ginseng* CA Meyer. *J Ginseng Res* 35(1):1–11
- Litholdo CG, Parker BL, Eamens AL, Larsen MR, Cordwell SJ, Waterhouse PM (2016) Proteomic identification of putative microRNA394 target genes in *Arabidopsis thaliana* identifies major latex protein family members critical for normal development. *Mol Cell Proteom* 15(6):2033–2047
- Livak KJ, Schmittgen TD (2012) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25(4):402–408
- Lynch M, Conery JS (2003) The origins of genome complexity. *Science* 302(5649):1401–1404
- Lytle BL, Song J, de la Cruz NB, Peterson FC, Johnson KA, Bingman CA, Phillips GN, Volkman BF (2009) Structures of two *Arabidopsis thaliana* major latex proteins represent novel helix-grip folds. *Proteins* 76(1):237–243
- Marković-Housley Z, Degano M, Lamba D, von Roepenack-Lahaye E, Clemens S, Susani M, Ferreira F, Scheiner O, Breiteneder H (2003) Crystal structure of a hypoallergenic isoform of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier. *J Mol Biol* 325(1):123–133
- Nam Y-W, Tichit L, Leperlier M, Cuerq B, Marty I, Lelièvre J-M (1999) Isolation and characterization of mRNAs differentially expressed during ripening of wild strawberry (*Fragaria vesca* L.) fruits. *Plant Mol Biol* 39(3):629–636

- Nawrot R, Lippmann R, Matros A, Musidlak O, Nowicki G, Mock H-P (2017) Proteomic comparison of *Chelidonium majus* L. latex in different phases of plant development. *Plant Physiol Biochem* 112:312–325
- Nessler CL, Burnett RJ (1992) Organization of the major latex protein gene family in opium poppy. *Plant Mol Biol* 20(4):749–752
- Nessler CL, Allen RD, Galewsky S (1985) Identification and characterization of latex-specific proteins in opium poppy. *Plant Physiol* 79(2):499–504
- Nessler CL, Kurz WG, Pelcher LE (1990) Isolation and analysis of the major latex protein genes of opium poppy. *Plant Mol Biol* 15(6):951–953
- Niu D-K, Jiang L (2013) Can ENCODE tell us how much junk DNA we carry in our genome? *Biochem Biophys Res Commun* 430(4):1340–1343
- Osmark P, Boyle B, Brisson N (1998) Sequential and structural homology between intracellular pathogenesis-related proteins and a group of latex proteins. *Plant Mol Biol* 38(6):1243–1246
- Qu Z-L, Wang H-Y, Xia G-X (2005) GhHb1: a nonsymbiotic hemoglobin gene of cotton responsive to infection by *Verticillium dahliae*. *Biochim Biophys Acta* 1730(2):103–113
- Radauer C, Lackner P, Breiteneder H (2008) The Bet v 1 fold: an ancient, versatile scaffold for binding of large, hydrophobic ligands. *BMC Evol Biol* 8(1):286
- Rebeiz M, Posakony JW (2004) GenePalette: a universal software tool for genome sequence visualization and analysis. *Dev Biol* 271(2):431–438
- Ruperti B, Bonghi C, Ziliotto F, Pagni S, Rasori A, Varotto S, Tonutti P, Giovannoni JJ, Ramina A (2002) Characterization of a major latex protein (*MLP*) gene down-regulated by ethylene during peach fruitlet abscission. *Plant Sci* 163(2):265–272
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4(4):406–425
- Siu-Ming Y, Wong Thomas KF, Zhao M, Ping N, Yang W, Yu L (2008) Promoter-sharing by different genes in human genome—*CPNE1* and *RBM12* gene pair as an example. *BMC Genom* 9(1):456
- Stanley Kim H, Yu Y, Sniesrud EC, Moy LP, Linford LD, Haas BJ, Nierman WC, Quackenbush J (2005) Transcriptional divergence of the duplicated oxidative stress-responsive genes in the *Arabidopsis* genome. *Plant J* 41(2):212–220
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10):2731–2739
- Thompson JD, Gibson T, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinform*. <https://doi.org/10.1002/0471250953.bi0203s00>
- Van LL, Rep M, Pieterse CM (2006) Significance of inducible defense-related proteins in infected plants. *Annu Rev Phytopathol* 44(1):135–162
- Venturini L, Ferrarini A, Zenoni S, Tornielli GB, Fasoli M, Dal SS, Minio A, Buson G, Tononi P, Zago ED (2013) De novo transcriptome characterization of *Vitis vinifera* cv. Corvina unveils varietal diversity. *BMC Genom* 14(1):41
- Verma V, Raju SC, Kapley A, Kalia VC, Dagainawala HF, Purohit HJ (2010) Evaluation of genetic and functional diversity of *Stenotrophomonas* isolates from diverse effluent treatment plants. *Bioresour Technol* 101(20):7744–7753
- Wang Y, Yang L, Chen X, Ye T, Zhong B, Liu R, Wu Y, Chan Z (2015) Major latex protein-like protein 43 (MLP43) functions as a positive regulator during abscisic acid responses and confers drought tolerance in *Arabidopsis thaliana*. *J Exp Bot* 67(1):421–434
- Weid M, Ziegler J, Kutchan TM (2004) The roles of latex and the vascular bundle in morphine biosynthesis in the opium poppy, *Papaver somniferum*. *Proc Natl Acad Sci USA* 101(38):13957–13962
- Wu JM, Wang Z-R, Hsieh T-C, Bruder JL, Zou J-G, Huang Y-Z (2001) Mechanism of cardioprotection by resveratrol, a phenolic antioxidant present in red wine. *Int J Mol Med* 8(1):3–17
- Xu W, Li R, Zhang N, Ma F, Jiao Y, Wang Z (2014) Transcriptome profiling of *Vitis amurensis*, an extremely cold-tolerant Chinese wild *Vitis* species, reveals candidate genes and events that potentially connected to cold stress. *Plant Mol Biol* 86(4–5):527–541
- Yang X, Tuskan GA (2006) Divergence of the Dof gene families in poplar, *Arabidopsis*, and rice suggests multiple modes of gene evolution after duplication. *Plant Physiol* 142(3):820–830
- Yang X, Kalluri UC, Jawdy S, Gunter LE, Yin T, Tschaplinski TJ, Weston DJ, Ranjan P, Tuskan GA (2008) The F-box gene family is expanded in herbaceous annual plants relative to woody perennial plants. *Plant Physiol* 148(3):1189–1200
- Yang C-L, Liang S, Wang H-Y, Han L-B, Wang F-X, Cheng H-Q, Wu X-M, Qu Z-L, Wu J-H, Xia G-X (2015) Cotton major latex protein 28 functions as a positive regulator of the ethylene responsive factor 6 in defense against *Verticillium dahliae*. *Mol Plant* 8(3):399–411
- Ye N, Jia L, Zhang J (2012) ABA signal in rice under stress conditions. *Rice* 5(1):1–9
- Ye CY, Li T, Yin H, Weston DJ, Tuskan GA, Tschaplinski TJ, Yang X (2013) Evolutionary analyses of non-family genes in plants. *Plant J* 73(5):788–797