

ARTICLE

Open Access

# The grapevine kinome: annotation, classification and expression patterns in developmental processes and stress responses

Kaikai Zhu<sup>1,2</sup>, Xiaolong Wang<sup>1,2</sup>, Jinyi Liu<sup>1</sup>, Jun Tang<sup>3</sup>, Qunkang Cheng<sup>4</sup>, Jin-Gui Chen<sup>5</sup> and Zong-Ming (Max) Cheng<sup>1,2</sup>

## Abstract

Protein kinases (PKs) have evolved as the largest family of molecular switches that regulate protein activities associated with almost all essential cellular functions. Only a fraction of plant PKs, however, have been functionally characterized even in model plant species. In the present study, the entire grapevine kinome was identified and annotated using the most recent version of the grapevine genome. A total of 1168 PK-encoding genes were identified and classified into 20 groups and 121 families, with the RLK-Pelle group being the largest, with 872 members. The 1168 kinase genes were unevenly distributed over all 19 chromosomes, and both tandem and segmental duplications contributed to the expansion of the grapevine kinome, especially of the RLK-Pelle group. *Ka/Ks* values indicated that most of the tandem and segmental duplication events were under purifying selection. The grapevine kinome families exhibited different expression patterns during plant development and in response to various stress treatments, with many being coexpressed. The comprehensive annotation of grapevine kinase genes, their patterns of expression and coexpression, and the related information facilitate a more complete understanding of the roles of various grapevine kinases in growth and development, responses to abiotic stress, and evolutionary history.

## Introduction

Reversible phosphorylation, catalyzed by protein kinases (PKs), is a type of post-translational modification that exists widely in both prokaryotic and eukaryotic organisms<sup>1</sup>. PKs constitute one of the largest gene families in eukaryotic genomes and are involved in numerous cellular and biological processes<sup>2,3</sup>. The majority of PKs transfer the  $\gamma$ -phosphate from adenosine triphosphate (ATP) to phosphorylate specific amino acids, mainly serine (S), threonine (T), or tyrosine (Y), in order to regulate the activity of their

target proteins<sup>4</sup>. Individual PK genes have been extensively analyzed in many species, including both prokaryotic and eukaryotic organisms and ranging from animals to plants in the latter group<sup>5</sup>. The first plant protein kinase was isolated from pea shoots (*Pisum sativum*) in 1973<sup>6</sup>, and the first plant PK DNA sequences were identified in bean (*Phaseolus vulgaris*) and rice (*Oryza sativa*) by using degenerate primers<sup>7</sup>. Since then, many plant PKs and PK gene families have been identified in plants, including *Arabidopsis*<sup>5,8,9</sup>, maize<sup>10</sup>, soybean<sup>3,11</sup>, tobacco<sup>12</sup>, and apple<sup>13</sup>.

The kinome, defined as the entire collection of kinase genes in a genome, was reported in *Arabidopsis* in 2004 as containing over 1000 PKs<sup>14</sup> and in rice as including ~1500 protein kinases<sup>15</sup>. Rice kinome meta-expression data provided evidence that rice PKs have functional roles in development and in stress tolerance<sup>16</sup>. More than 2000

Correspondence: Z.-M. Cheng (zmc@njau.edu.cn) (zcheng@utk.edu)

<sup>1</sup>College of Horticulture, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China

<sup>2</sup>Department of Plant Sciences, University of Tennessee, Knoxville, TN 37996, USA

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

© The Author(s) 2018



**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

kinase genes make up the soybean kinome, which is approximately twice as many as that in *Arabidopsis* and four times the size of the human kinome<sup>3</sup>. In general, however, the number of protein kinases in plants is much greater than those in humans and animals<sup>17,18</sup>. The expansion of protein kinases in plants may be the result of recent duplication events, whole-genome duplication, segmental duplication, tandem duplication, or all three<sup>19</sup>.

Protein kinases generally possess a catalytic domain, which is composed of 250–300 amino acid residues<sup>2</sup>. Classification of PKs was initially based on a phylogenetic analysis of the kinase domain<sup>20</sup>. Lehti-Shiu and Shiu<sup>2</sup> classified the PKs of 25 sequenced plant species into 9 major groups including AGC (PKA-PKG-PKC), CAMK (calcium- and calmodulin-regulated kinase), CK1 (casein kinase 1), CMGC (cyclin-dependent kinases, mitogen-activated protein kinases, glycogen synthase kinases and cyclin-dependent like kinases), STE (sterility), TK (tyrosine kinase), TKL (tyrosine kinase-like kinase), and plant-specific.

Grape (*Vitis vinifera* L.) is one of the most popular and economically important fruits in the world<sup>21,22</sup>. Grape production, similar to that of other crops, however, must address abiotic and biotic stresses, which cause reductions in yield and fruit quality. The publication of the grapevine genome<sup>23</sup> has provided tremendous new opportunities for comprehensive transcriptomic and proteomic studies, as well as other genetic and molecular studies. Information on several grapevine kinase gene families has been reported, including mitogen-activated protein kinase (MAPK)<sup>24</sup>, Ca<sup>2+</sup>-dependent protein kinase (CDPK)<sup>25</sup>, and sucrose non-fermenting 1-related protein kinase 2 (SnRK2)<sup>26</sup>. In addition, a preliminary analysis of the entire superfamily of grapevine protein kinases was first reported in 2012; 877 grapevine kinase genes were identified and classified into 118 families<sup>2</sup>. In the current study, 1168 putative PK coding genes were identified in the most recent grapevine genome database and classified into groups and families. We further analyzed the distribution of kinases in the grapevine genome, intron numbers, subcellular location prediction, expansion mechanisms, and gene expression patterns in different tissues during development and in response to a variety of stress treatments. In addition, a coexpression network was constructed. This study provides a comprehensive overview of the entire grapevine kinome and the potential roles of different groups of kinases in grapevine growth and stress response and adaptation.

## Materials and methods

### Computational identification and classification of grapevine protein kinases

Grapevine (*Vitis vinifera*) predicted proteins (Version 2.1) were downloaded from the Grape Genome Biotechnology Center (CRIBI, <http://genomes.cribi.unipd.it>)

and used to identify the entire kinome of grapevine. Files containing Hidden Markov Models (HMMs) of the Pkinase clan (Pkinase (PF00069) and Pkinase\_Tyr (PF07714)) were downloaded from Pfam (<http://pfam.xfam.org/>) and used to identify putative protein kinases<sup>27</sup> using HMMER v. 3.1b1 software with an E-value cutoff of <1.0<sup>28</sup>. The presence of a kinase domain in each of the candidate PK genes was verified using Pfam and SMART (<http://smart.embl-heidelberg.de/>). Subsequently, each sequence containing the protein kinase domain was extracted, and redundant sequences were deleted. Only the longest peptide sequence of each gene was used in cases where an identified gene contained more than one predicted isoform. This approach resulted in the identification of a total of 1358 grapevine protein kinases. An identified sequence was considered a putative protein kinase if the kinase domain alignment covered at least 50% of the Pfam domain model<sup>2,3</sup>. This selection criterion resulted in the identification of a final total of 1168 proteins containing at least one kinase domain (Supplementary Table S1) and identified as grapevine protein kinases. The identified grapevine PKs were then classified into different groups and families using HMMs of the different kinase families developed from the PK sequences of 25 plant species<sup>2</sup>.

### Multiple sequence alignment and phylogenetic analysis

The kinase domain sequences of all 1168 grapevine PKs were aligned using MUSCLE, and a phylogenetic tree was constructed using the Neighbor-Joining (NJ) method in MEGA 6.06<sup>29</sup>. The parameter pairwise deletion and the Poisson model were used with 1000 bootstrap replicates.

### Intron numbers and chromosomal locations

The chromosomal positions of grapevine PK genes were collected from CRIBI. Intron numbers of all of the grapevine kinase genes were obtained from the GFF file from CRIBI.

### Protein properties and subcellular localization prediction

The molecular weight (MW) and isoelectric point (pI) for all of the identified grapevine PK protein sequences were predicted using the ExPASy server tool ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/))<sup>30</sup>. The predicted subcellular localizations of the grapevine PKs were determined using the online tool WoLF PSORT (<http://www.genscript.com/wolf-psort.html>)<sup>31</sup>.

### Identification of tandem and segmental duplication events

Tandem duplications of grapevine kinase genes were predicted by determining their physical location on grapevine chromosomes. An array of two or more genes located on the same chromosome and separated by five or fewer genes in a 100-kb region was considered to represent a tandem duplication event<sup>32</sup>. An illustration of the

chromosomal locations of tandemly duplicated genes was constructed using Mapchart 2.30 (<http://www.wur.nl/en/show/Mapchart-2.30.htm>). The segmental duplication events of grapevine kinases were determined using the MCScanX (Multiple Collinearity Scan, <http://chibba.pgml.uga.edu/mcscan2/>) package<sup>33</sup>.

#### GO analysis of tandemly duplicated PK genes

The Gene Ontology (GO) term ID of each tandemly duplicated PK gene in grapevine was retrieved from the CRIBI database (<http://genomes.cribi.unipd.it/DATA/V2/annotation/bl2go.annot.txt>). The annotations of the GO term IDs were collected from the Gene Ontology Consortium (<http://www.geneontology.org>).

#### Calculations of synonymous substitutions (*Ks*) and non-synonymous substitutions (*Ka*) and selection modes

The coding sequences (CDS) of the tandemly and segmentally duplicated grapevine PK genes in each family were aligned using ClustalW 2.0<sup>34</sup> in order to detect the selection modes of the tandem and segmental duplication events. The *Ka*, *Ks*, and ratio of *Ka* to *Ks* (*Ka/Ks*) were calculated using MEGA 6.06<sup>29</sup>. The approximate divergence times (*T*) of the duplication events were calculated using the equation  $T = Ks / 2\lambda \times 10^{-6}$  MYA (million years ago), in which the mean *Ks* rate  $\lambda = 6.5 \times 10^{-9}$  for grapevine<sup>35</sup>.

#### Microarray data analysis

Genome-wide expression data for grapevine from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>) at the NCBI website (<https://www.ncbi.nlm.nih.gov/>) were analyzed for the expression patterns of grapevine PK gene families in different tissues and developmental stages and in response to different treatments. The GSE36128 dataset provided gene expression profiling of 54 diverse grapevine samples covering different growth stages of most grapevine organs<sup>21</sup>. The expression data of 1,109 available grapevine kinase genes were collected and normalized. The gene expression data of 231 available kinase genes in response to salt, polyethylene glycol (PEG), cold, drought, heat and abscisic acid (ABA) treatments were downloaded from four selected Affymetrix datasets, GSE31594, GSE31677, GSE31675, and GSE31662. These datasets were used to analyze the expression of grapevine PKs in response to different stress treatments. The expression data on the response of *Arabidopsis* PK genes to drought (GSE65414) and salt (GSE71855) were also retrieved to compare gene expression between the grapevine kinome and the *Arabidopsis* kinome. Expression data were analyzed using the GEO2R online tool (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>). To analyze the expression data at the family level, the normalized gene expression data of all members in each family were averaged. Since

expression data for only 1109 of the 1168 identified grapevine kinase genes was available in 54 different tissues and the responses of only 231 of the 1168 grapevine PKs to different stress treatments were available, only families where at least 33% of the family gene members were represented in the expression data were considered<sup>3</sup>. Heatmaps representing the expression data of 120 families in different tissues at different developmental stages and the responses of 56 families to different stress treatments were constructed using the R package (<https://www.r-project.org/>) pheatmap.

#### Coexpression network of grapevine PK families

A coexpression network of 56 grapevine PK families in response to six different treatments was constructed based on Pearson correlation coefficients (PCCs) to better understand the topological relationships between stress-responsive grapevine PK families. All of the available gene expression data in each family were averaged, and the PCC was calculated between each pair of families using the method described by Tang et al<sup>36</sup>. All of the grapevine PK family pairs with PCCs at 0.05 significance level (*p*-value) were extracted and used to illustrate a coexpression network using Cytoscape 3.3 (<http://www.cytoscape.org>).

#### Plant materials, growth conditions and treatments

In vitro-grown grapevine plants (*Vitis vinifera*, Pinot noir 40024, the sequenced genotype) were maintained on half-strength Murashige and Skoog medium containing 0.3 mg/L indole 3-butyric acid (IBA) and grown at 25 °C in the culture room under a photoperiod of 16/8 h and a light intensity of 100  $\mu\text{mol m}^{-2} \text{s}^{-1}$ . Five-week-old tissue-cultured grapevine plants were transplanted into pots and acclimatized in a growth chamber with 16 h light at 24 °C/ 8 h dark at 22 °C and 70–80% relative humidity. Plants grown in pots were well watered for the first four weeks, and then, water was withheld to impose drought treatment. Grapevine leaf samples were collected at 0 (control), 2, 4, and 8 days after water was withheld. For salt treatment, in vitro-grown 5-week-old plants were irrigated with 200 mM NaCl directly and grown at 25 °C under a photoperiod of 16/8 h. Leaves were collected at 0 (control), 6, 12, and 24 h and used in real-time quantitative reverse transcription PCR (RT-PCR) analysis. All the collected samples were immediately frozen in liquid nitrogen and stored at −70 °C for further analysis. Three biological replicates were used for each treatment or control.

#### RNA isolation and real-time quantitative RT-PCR analysis

Total RNA was extracted from grapevine leaves using the Plant Total RNA Isolation Kit (Foregene, Chengdu, China) according to the manufacturer's instructions. The genomic DNA was removed from the sample using DNase I (RNase-

free DNase set, Qiagen, Hilden, Germany), and RNA concentration and quality were measured by determining the optical density (OD) ratios at 260 nm and 280 nm with a One Drop OD-1000 spectrophotometer (Thermo Scientific, Wilmington, DE). Then, first-strand cDNA was synthesized from 1 µg total RNA using Prime Script RT Reagent Kit (Takara, Japan) according to the manufacturer's instructions. Specific primers for the grapevine PK genes were designed using the IDT PrimerQuest tool (<https://www.idtdna.com/PrimerQuest/Home/Index>). The actin gene (*VIT\_212s0178g00200*) was used as a reference gene for data normalization; we have confirmed this to be an appropriate reference gene<sup>24,26</sup>. The qRT-PCR was conducted using a LightCycler 480 SYBR Green I Master (Roche, Mannheim, Germany) on a Roche 480 Real-Time PCR System (Roche, Mannheim, Germany). The qRT-PCR was performed according to the manufacturer's protocol, and the total volume of each reaction was 15 µl. The PCR conditions used were as follows: 95 °C 10 min, 40 cycles of 10 s at 95 °C, 20 s at 58 °C, and 10 s at 72 °C. At the end, a melting curve was generated from 65 to 97 °C. The relative gene expression levels in stressed plants relative to control were shown using the  $\Delta\Delta C_T$  method<sup>11</sup>. Values were means  $\pm$  SE calculated from the three biological replicates.

### RNA-Seq analyses of grapevine PK genes

Three biological replicates of grapevine leaf samples under the aforementioned drought treatment were collected and used for an RNA-Seq experiment. A total amount of 1 µg RNA per sample was used as input material for the RNA library preparations. Sequencing libraries were generated using NEBNext Ultra™ RNA Library Prep Kit for Illumina (NEB, USA) following manufacturer's recommendations. The transcript abundance of the grapevine kinase genes was calculated as fragments per kilobase of exon model per million mapped reads (FPKM). The RNA-Seq raw data have been uploaded to the Sequence Read Archive of NCBI (<https://www.ncbi.nlm.nih.gov/sra/>) with accession number PRJNA433817.

### Statistical analysis

Statistical analysis was performed using either *t*-test ( $P < 0.05$ ) or Duncan's multiple range test ( $P < 0.05$ ) where appropriate using IBM SPSS Statistics v24 software (SPSS, Inc., USA).

## Results

### Genome-wide identification and classification of grapevine protein kinases

The putative PK sequences in the grapevine genome database were identified by HMMER and aligned to the Pfam kinase domain. A total of 1,168 typical putative grapevine PK genes (kinase domain aligned with at least 50% of the kinase domain models; Supplementary Table S1) and 190 atypical

PKs (Supplementary Table S2) were identified. The 1,168 typical PKs were classified (Fig. 1; Supplementary Table S3) and further examined using a phylogenetic analysis (Supplementary Fig. S1; Supplementary Fig. S2). The results of the phylogenetic analysis indicated that only 19 of the 1168 genes exhibited a classification that was different from the HMM search result. These 19 genes had low E-values and did not cluster with any other known families. Thus, they were placed in an unclassified cluster. The remaining 1149 grapevine PKs were classified into 20 groups, which included AGC, Aur, BUB, CAMK, CK1, CMGC, IRE1, NAK, NEK, PEK\_GCN2, RLK-Pelle, SCY1, STE, TKL, TLK, TTK, ULK, WEE, WNK, and plant-specific (Fig. 1). Some of the larger groups were further classified into different families within each group<sup>2,3</sup>. The overall classification resulted in 20 groups and 121 families (Fig. 1; Supplementary Table S3). The RLK-Pelle group was the largest group, accounting for 74.66% (872/1168) of the grapevine PK genes.

In comparison to the classifications of *Arabidopsis* and rice and to the grapevine kinomes reported previously<sup>2</sup>, 35 grapevine families contained only 1 member, including AGC\_PKA-PKG, CMGC\_CDK-CCRK, CMGC\_CDK-CDK8, CMGC\_PL-The, RLK-Pelle\_RLCK-II, and TTK (Fig. 1). Most of these grapevine PKs were also conserved in the rice and *Arabidopsis* kinomes. Surprisingly, the RLK-Pelle\_LRK10L-2 and RLK-Pelle\_LRR-XI-1 families contained many more genes in grapevine than in *Arabidopsis* and rice, indicating that these two families have expanded greatly in grapevine. However, the RLK-Pelle\_LRK10L-2 family (8) was much smaller<sup>2</sup>. Five families were newly found in the grapevine kinome, including BUB, CMGC\_CDK-CCRK, RLK-Pelle\_LRR-Xb-2, RLK-Pelle\_LRR-XV, and RLK-Pelle\_RLCK-IXa (Fig. 1). The monocot-specific kinase families (GMGC\_CDKL-Os, 18 members) were not found in either *Arabidopsis* or grapevine<sup>3</sup>. RLK-Pelle\_RLCK-IXa contained only one member in grapevine and was absent in rice and *Arabidopsis*.

### The kinase domains, intron numbers, protein properties, and predicted subcellular localizations of grapevine PKs

The sequence of the kinase domain, intron number, and protein properties, including molecular weight (MW) and isoelectric point (pI) (Supplementary Table S3), were characterized for each of the 1168 grapevine PKs. Most (95.21%) grapevine protein kinases possess only one catalytic domain, 54 possess two catalytic domains, and 2 possess three kinase domains (*VIT\_209s0002g02840* and *VIT\_209s0002g02840*) (Supplementary Table S4). The 56 grapevine protein kinases with two or more kinase domains were distributed in 18 different families. The average intron number in the 1,168 kinase genes was 5, with 194 PK genes being intronless genes, of which 173 were RLK-Pelle genes. The number of introns per gene varied from 0 to 49 (*VIT\_212s0055g00590*). Only 147 (12.59%) of the grapevine kinase genes contained more

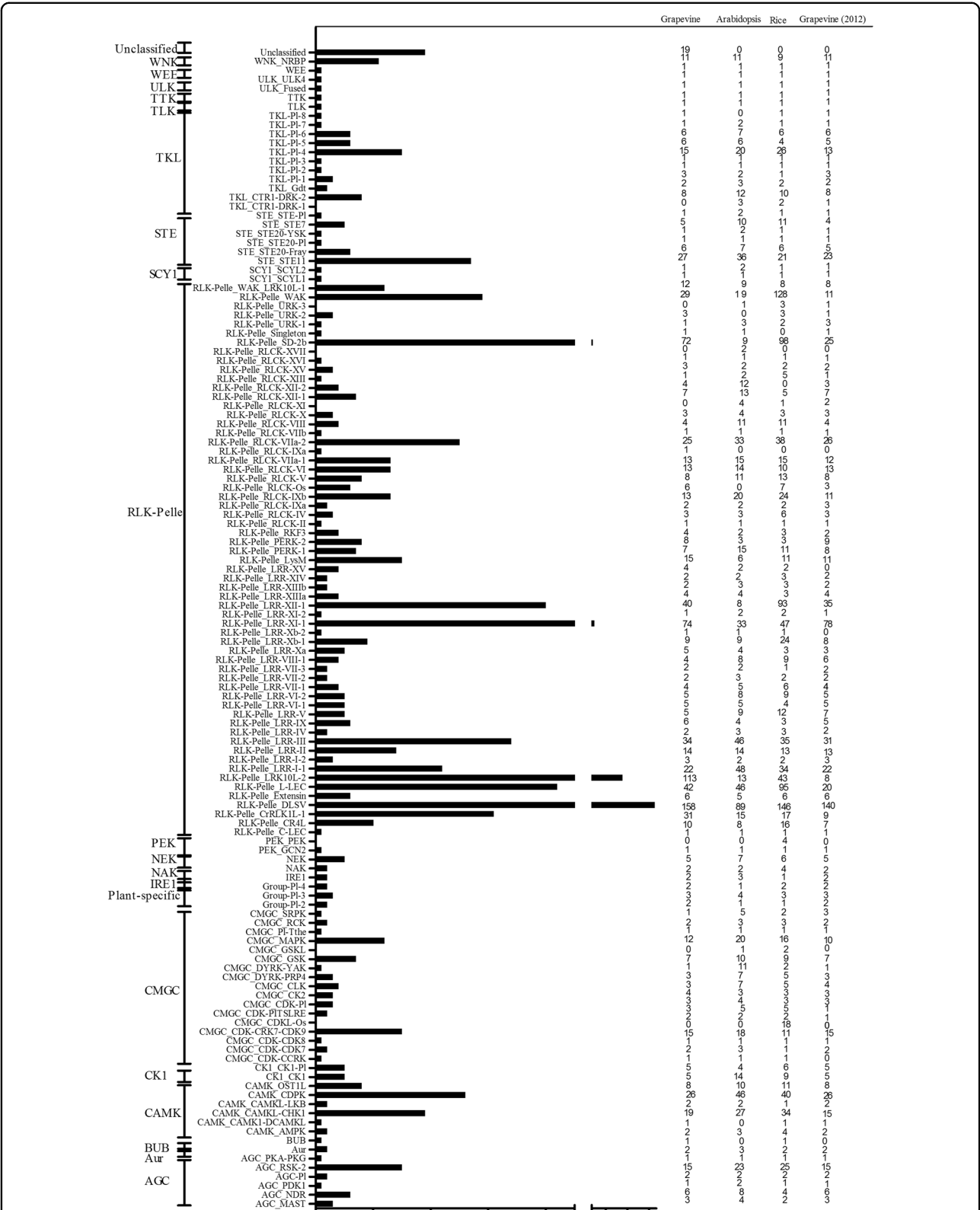


**Fig. 1** Comparison of grapevine protein kinase families with those in *Arabidopsis* (eudicot species) and rice (monocot species). The grapevine PKs were classified into 20 groups and 121 families

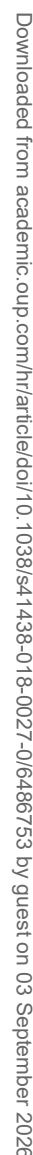
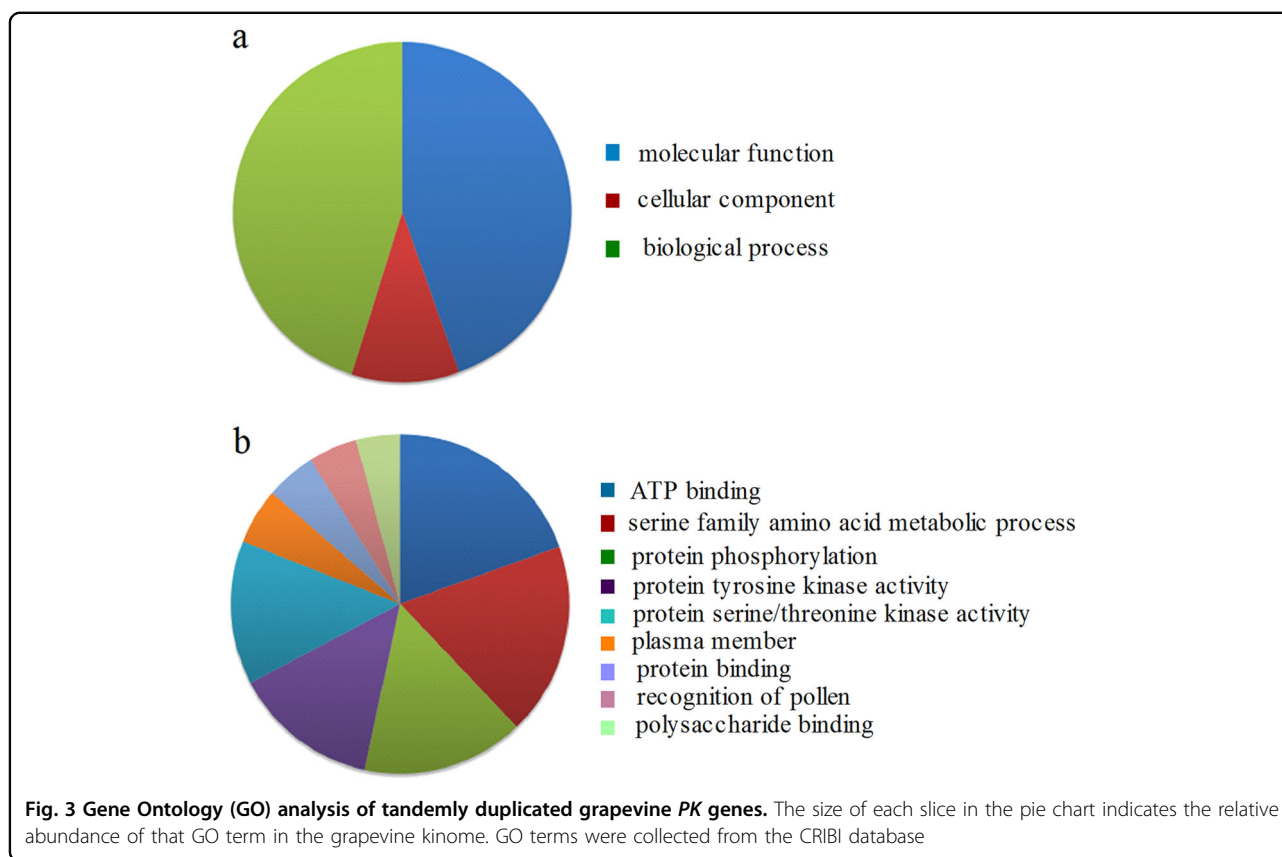


**Fig. 2 Chromosomal locations of tandemly duplicated kinase genes in grapevine.** A total of 395 tandemly duplicated kinase genes representing 19 families were distributed unevenly among 18 chromosomes. Gene IDs and related family names are presented to the right of each chromosome, and related information on gene location is listed on the left.



than 10 introns, and 26 (2.23%) contained more than 20 introns. The pIs of the grapevine PKs varied from 4.48 to 10.19, with MWs ranging from 16,243.53 to 230,303.18 daltons (Supplementary Table S3). In general, grapevine PK genes within the same families shared similar numbers of introns, as well as similar pIs and MWs.

The subcellular localizations of the grapevine PKs were determined using WoLF PSORT software (Supplementary Table S3). Most of the PKs (297) were predicted to localize to the plasma membrane, whereas 223, 225, and 248 were predicted to reside in the cytoplasm, nucleus, and chloroplast, respectively. Several families were also predicted to have the same subcellular locations for all members. For example, AGC\_MAST, CK1\_CK1, CMGC\_CDK-PITSLRE, and TKL-PI-6 family members were all predicted to localize to the nucleus. Members of the Aur, CAMK\_AMPK, CMGC\_CDK-PI, and Group-PI-4 families were all predicted to be cytoplasmic. Many other PKs within the same families, however, were predicted to localize to different cellular compartments.

#### Chromosomal locations and tandem duplications of grapevine PK genes

Almost all of the PK genes (1063/1168) could be mapped to the 19 grapevine chromosomes (Supplementary Table S3), but they are not evenly distributed among the 19

chromosomes. Chromosome 16 hosts the largest number (107) of PK genes, whereas chromosome 2 has the fewest (23). Grapevine PK members in the same family are generally distributed on different chromosomes. Interestingly, most RLK-Pelle\_LRK10L-2 gene family members are located on chromosome 16, and the RLK-Pelle\_LRR-I-1 gene family members are mainly distributed on chromosome 9. All of the RLK-Pelle\_RLCK-XII-2 gene family members are located on chromosome 18.

Grapevine PK genes are tandemly arranged on chromosomes based on their physical distribution. It appears that 395 PK genes were generated by tandem duplication (Fig. 2; Supplementary Table S5). These tandem duplication events occurred in 19 families that were unevenly distributed across all of the chromosomes with the exception of chromosome 2. Chromosome 16 possesses the greatest number (82) of PK genes, and 78 of these are RLK-Pelle\_LRK10L-2 genes. Tandem duplication events were responsible for more than 60% of the expansion in four families, RLK-Pelle\_LRK10L-2 (76.11%), RLK-Pelle\_LRR-I-1 (77.27%), RLK-Pelle\_RLCK-Os (66.67%), and RLK-Pelle\_RLCK-XII-2 (75%), all of which belong to the RLK-Pelle group.

Previous studies have indicated that genes that are duplicated by tandem duplication events may play distinct roles in plant development and response to abiotic and

biotic stresses<sup>37</sup>. Therefore, the GO annotations for 395 tandemly duplicated grapevine PK genes were examined (Fig. 3). The three main GO categories are biological process, cellular component, and molecular function. The majority of grapevine PKs that arose by tandem duplication were enriched in two of the three GO categories, namely, biological process and molecular function, with only approximately 10% of tandemly duplicated PK genes enriched in the GO category of cellular component (Fig. 3a). Functional GO terms for the 395 tandemly duplicated grapevine PK genes were also assessed. The top two terms were ATP binding (GO:0005524) and serine family amino acid metabolic process (GO:0009069; Fig. 3b). These results suggest that these PKs are associated with multiple biological processes, as well as signaling and response to stress.

#### ***Ka* and *Ks* calculations and selection modes of tandem and segmental duplication events**

In addition to tandem duplication events, which were the main contributors to the expansion of grapevine PKs, segmental duplication events played a role. Collectively, tandem and segmental duplication events were identified (Supplementary Table S6). These segmental duplication events include 231 PK genes from 47 families, located on all 19 chromosomes.

To determine the selection types and divergence dates of the tandem and segmental duplication events associated with the grapevine PK genes, the synonymous (*Ks*) and non-synonymous substitutions (*Ka*) between the gene pairs were determined. The substitution ratio of *Ka* to *Ks* was calculated in order to evaluate the selection model of the gene pairs.  $Ka/Ks < 1$  indicates purifying selection on the gene pairs,  $Ka/Ks = 1$  indicates neutral selection, and  $Ka/Ks > 1$  indicates that the gene pairs are under positive selection<sup>38</sup>. A summary of the *Ka/Ks* ratios for the 585 tandem and 138 segmental duplication events is shown in Supplementary Table S6. The *Ka/Ks* values in 101 of the segmentally duplicated gene pairs were less than 1. Only 37 of the gene pairs were under positive selection, and all of these were members of the RLK/Pelle group. Similar results were observed for the PK genes associated with tandem duplication events. In this case, the *Ka/Ks* values of 541 of the 585 gene pairs were less than 1. Collectively, these data indicate that the majority of tandemly and segmentally duplicated genes in the grapevine kinome were subjected to purifying selection (Supplementary Fig. S3; Supplementary Table S6). The divergence dates of the tandemly and segmentally duplicated gene pairs occurred from 0.30 to 387.18 and 0.35 to 374.66 MYA, respectively (Supplementary Table S6).

#### **Gene expression patterns of grapevine PKs in different tissues**

Protein kinases have been reported to play critical roles in plant development<sup>39,40</sup>. Therefore, the expression

profiling data of 54 different grapevine tissues and organs during different developmental stages (accession #GSE36128) were analyzed to obtain insight into the temporal transcription patterns of grapevine PK genes during development. Expression data for 1,109 of the 1,168 grapevine PK genes were available, and the expression of these genes in the acquired datasets was analyzed at the PK family level (Supplementary Table S7). Cluster analysis of the microarray data reflecting the expression of 120 PK families was calculated in the 54 different tissues (Fig. 4). The results indicated that different families exhibit different expression patterns during plant development. Most families in the AGC, CAMK, CMGC, SCY1, STE, and TKL groups exhibited high expression levels in most tissues, suggesting that these families play key roles in developmental processes. In contrast, Aur, BUB, ULK, WEE and most families in the RLK-Pelle groups exhibited very low expression levels in most of the different tissues (Fig. 4). Although the RLK-Pelle group contains 58 families, only RLK-Pelle\_C-LEC, RLK-Pelle\_LRR-VII-2, RLK-Pelle\_LRR-Xb-2, and RLK-Pelle\_RLCK-II exhibited high expression levels in nearly all 54 tissues. Other families had high levels of expression in specific tissues. For example, RLK-Pelle\_RLCK-VIIb exhibited high levels of expression in seed-PFS (post fruit set seed), and RLK-Pelle\_PERK-1, RLK-Pelle\_RLCK-IXb, RLK-Pelle\_RLCK-VIIa-1, and RLK-Pelle\_RLCK-VIII had high expression levels in flower, pollen, and stamens (Fig. 4).

#### **Microarray analyses and the related coexpression network of grapevine PK families in response to different stress treatments**

PKs play key roles in plant responses to various environmental stresses<sup>3,24,37</sup>. Therefore, the publicly available grapevine gene expression data from NCBI for responses to salt, PEG, cold, drought, heat, and ABA treatments (accessions GSE31594, GSE31677, GSE31675, and GSE31662) (Supplementary Table S8) were analyzed. Expression data covering 231 grapevine PK genes from 56 families indicated that half of the families exhibited significant up-regulation or downregulation in response to specific stresses (Fig. 5). Based on a cluster analysis, we found that some of these families were induced by multiple similar stress treatments. For example, AGC\_PDK1 and RLK-Pelle\_LRR-II were induced by salt, PEG, and heat. CAMK\_CAMKL-CHK1, IRE, RLK-Pelle\_LRR-VI-2, RLK-Pelle\_RLCK-IV, and TKL-PI-4 were upregulated in response to salt, PEG, and drought. Several PK families, such as Aur, RLK-Pelle\_LRR-XIIIb, RLK-Pelle\_LRR-IV, and RLK-Pelle\_LRR-VII-3, were downregulated in response to salt, drought, and PEG stress. Other families exhibited contrasting expression patterns in response to different treatments. CMGC\_CDK-PI, RLK-Pelle\_LRR-

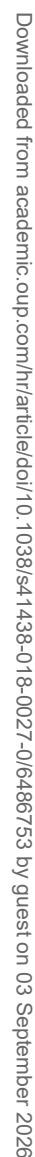


**Fig. 4 A heatmap illustrating the expression data of 120 different grapevine kinase families in 54 different grapevine tissues at various stages of development.** The heatmap was generated using the R package pheatmap. The color scale represents expression levels, with red indicating high expression levels and blue indicating low levels

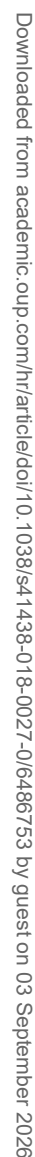
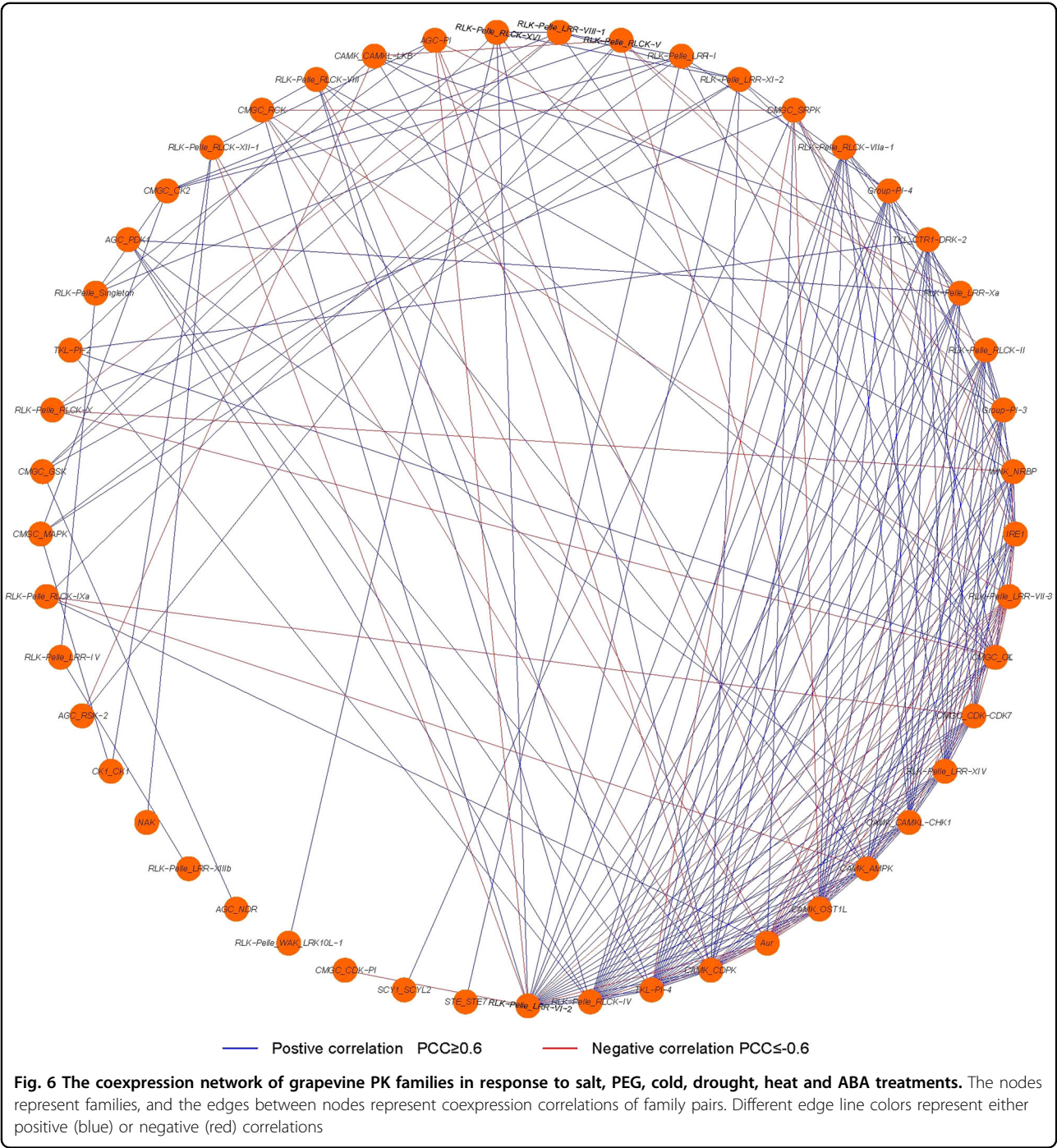


**Fig. 5 A heatmap of the expression data of 56 different grapevine kinase families in response to salt, PEG, cold, drought, heat, and ABA.** The heatmap was generated using the R package pheatmap. The color scale represents expression levels, with red indicating high expression levels and blue indicating low levels



XIIIb, and RLK-Pelle\_RLCK-IXa were downregulated in response to salt, PEG, and drought treatments but upregulated in response to heat stress.

A coexpression network was constructed in order to discover the mutual relationships between grapevine PK families by analyzing a large number of grapevine PKs with similar expression patterns under different conditions<sup>41,42</sup>. A coexpression network of grapevine families was constructed using the expression data described

above to gain additional insight into the relationships and potential interactions between different grapevine PK families in response to six different treatments. The resulting coexpression network (Fig. 6) contained 49 nodes (grapevine PK families) and 222 edges (coexpression family pairs). Among 56 grapevine PK families, 49 were linked to each other through 222 edges, representing PCCs between coexpression events. Each node harbored a different number of regulatory edges, varying from 1 to

21. The TKL-PI-4 family contained the maximum number of edges. Another five families, RLK-Pelle\_RLCK-IV, WNK\_NRBP, RLK-Pelle\_LRR-VI-2, RLK-Pelle\_LRR-VII-3, and RLK-Pelle\_LRR-XIV, contained more than 10 edges, suggesting that these families may play central roles in the response of grapevine to different stress treatments. Among the 222 regulatory edges, 171 had significantly positive correlations, and the remaining 51 coexpression events had significantly negative correlations. A list of 20 kinase genes that were significantly induced ( $>two\text{-fold}$ ,  $P < 0.05$ ) by at least three different stresses (Supplementary Table S9) was constructed based on the expression and coexpression data. These 20 PK genes belonged to 4 different groups and 14 families. These PKs may play key roles in stress signaling pathways in grapevine and merit further investigation.

#### RNA-Seq and qRT-PCR expression analysis of grapevine PK genes under drought and salt stresses

In addition to the public microarray data analysis, RNA-Seq analysis was also performed to analyze the functions of grapevine PKs under drought treatment. In total, 969 available grapevine PK genes showed their expression patterns in response to drought with FPKM values (Supplementary Table S10). We found that a large fraction (571/969) of the PK genes showed low FPKM values of less than 10 in all samples (Supplementary Fig. S4; Supplementary Table S10). Different PK genes showed various expression patterns. Some grapevine PK genes could be induced by drought treatment for at least one time point. However, some PK genes were downregulated under drought stress.

Twelve PK genes that were induced by drought in both the microarray (Supplementary Table S8) and RNA-Seq (Supplementary Table S10) analyses were selected for further confirmation using qRT-PCR after imposing drought stress (Fig. 7a). Specific primers for the 12 grapevine PK genes used in this study are listed in Supplementary Table S11. Interestingly, all these 12 PK genes were significantly induced by drought treatment at different time points. Most of them were highly upregulated at the 8-day time point only, with the exception of *VIT\_201s0011g03010*, which was significantly induced only after 2 days of drought application. The expression patterns of most of the selected genes were similar in the RNA-Seq and qRT-PCR analyses, also indicating that the RNA-Seq data were reliable (Fig. 7a). We further analyzed the expression patterns of these 12 PK genes when plants were exposed to 200 mM NaCl (Fig. 7b). The 12 genes showed various patterns under salt stress. For instance, four kinase genes including *VIT\_208s0058g01040*, *VIT\_204s0008g05770*, *VIT\_205s0020g00830*, and *VIT\_208s0058g01180* showed higher expression levels than those of the control after exposure to NaCl for 12 or 24 h and were also upregulated under drought stress. Only one gene (*VIT\_206s0004g03670*) had no

significant changes under salt treatment. The remaining seven PK genes were downregulated at different time points. Overall, understanding the expression patterns of the grapevine PK genes might be valuable for improving the stress resistance of grapevine.

#### Discussion

Plant kinases play a central role in signal transduction during plant growth and development and in plant responses to biotic and abiotic stresses. Comprehensive information on the grapevine kinome could provide further insight into the potential functions of PKs in grapevine.

#### The grapevine genome possesses a large kinome

It has been reported that ~4% of the genes in a plant genome are kinase genes<sup>43</sup>. The grapevine genome contains 1168 putative grapevine kinase genes (Supplementary Table S1), accounting for ~3.7% of the total number of predicted genes. The number of grapevine PKs (1,168) is similar to the numbers in *Arabidopsis* (1008), maize (1241), and tomato (1436), but significantly more than the number previously reported (877) in grapevine<sup>2,40,44</sup>. It is, however, much smaller than the number of PKs in soybean (2,166)<sup>3</sup> and larger than the kinomes in human<sup>18</sup>, mice<sup>17</sup>, oysters<sup>45</sup>, and the fungus *Wolfiporia cocos*<sup>46</sup>. The number of grapevine kinase families (121) (Fig. 1; Supplementary Table S3, Supplementary Fig. S1; Supplementary Fig. S2) is also similar to the numbers in *Arabidopsis* (119), rice (123), and soybean (122)<sup>2,3</sup>. The RLK-Pelle group is the largest kinase group in grapevine, accounting for 74.66% of the kinome, which is also similar to the proportion in other plant species<sup>3,37,40</sup>. Some grapevine PK families, however, had very few members. Members of these families may be involved in more basic cellular processes and less in processes associated with responses to stress<sup>43</sup>. Most grapevine PKs possessed only one kinase catalytic domain. Fifty-six of the grapevine PKs, however, contained two or three kinase domains (Supplementary Table S4). The grapevine kinases with two or three kinase domains may be required for specific substrates<sup>3</sup>. As reported for maize<sup>40</sup>, the intron number, pI, and MW of PKs within the same family were generally similar (Supplementary Table S3). Grapevine PK genes were unevenly distributed on all 19 chromosomes. However, many were tandemly arranged on a specific chromosome (Supplementary Table S3).

#### Tandem and segmental duplications contributed greatly to the expansion of the grapevine kinome

In addition to whole-genome duplications, tandem and segmental duplication events have played major roles in the evolutionary expansion of gene families<sup>19</sup>. The existence of duplicated genes can also foster the acquisition of

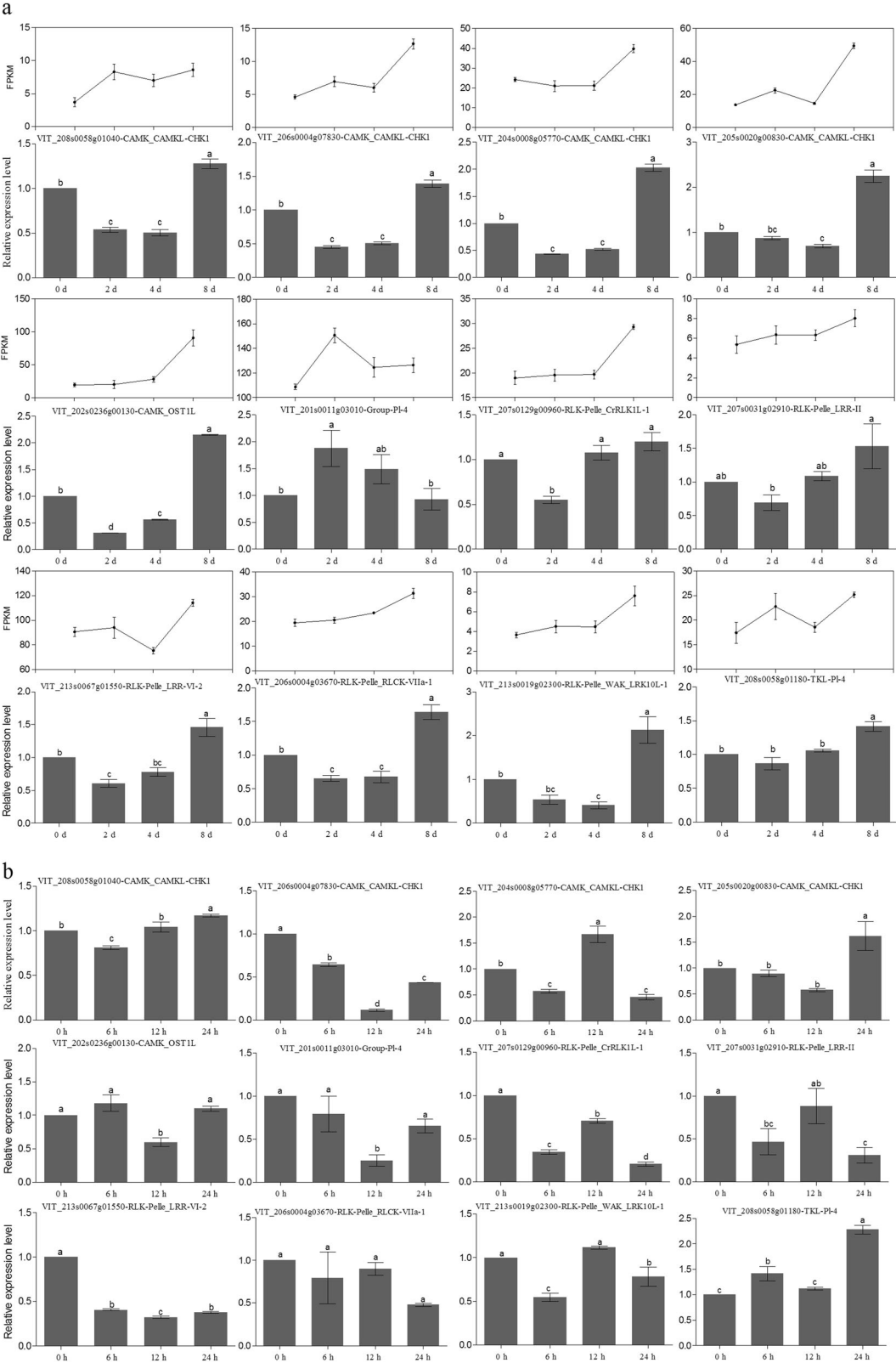


**Fig. 7** (See legend on next page.)

(see figure on previous page)

**Fig. 7 The expression patterns of grapevine PK genes under drought and salt stresses.** **a** qRT-PCR expression data of 12 selected PK genes subjected to drought stress by withholding water for 0 (control), 2, 4, and 8 d after full water saturation. The RNA-Seq data of related grapevine PK genes are shown as FPKM values. **b** Quantification of expression levels of 12 grapevine PK genes in response to salt stress using qRT-PCR. Plants were subjected to salt stress by irrigating with 200 mM NaCl, and leaves were collected at 0 (control), 6, 12, and 24 h post salt treatment. The actin gene (*VIT\_212s0178g00200*) was used as a reference gene to normalize gene expression levels. Values are the means  $\pm$  SE of three replicates for each, and bars with different letters are significantly different at  $P < 0.05$  according to Duncan's multiple range test

expanded functions for the new genes<sup>47</sup>. Tandem duplications tend to initiate changes in gene structure and function more quickly than other mechanisms of duplication<sup>37</sup>. Approximately one-third (395) of the grapevine kinome arose by tandem duplication, and 382 of the tandem duplication events occurred in members of the RLK-Pelle group of PKs (Fig. 2; Supplementary Table S5). The GO functional analysis revealed that the group of tandemly duplicated kinase genes as a whole was enriched in the GO categories “molecular function” and “biological process” (Fig. 3), suggesting that tandemly duplicated genes as a whole may play a central role in signaling pathways involved in plant development and responses to stress<sup>3</sup>.

Nearly two-thirds of the detected segmentally duplicated kinase genes (153/231) were also RLK-Pelle members. Additionally, based on the analysis of *Ka/Ks* ratios, it appears that the tandem and segmental duplication events in the grapevine kinome were subjected to purifying selection (Supplementary Table S6), which is also true for the soybean kinome<sup>3</sup>. The *Arabidopsis* kinome contains over 600 protein kinases that are RLK-Pelle group members, and the expansion of this group in *Arabidopsis* has also been largely attributed to tandem and segmental duplication events<sup>14,48</sup>. This type of expansion has also been found to be true for the kinomes of maize<sup>40</sup> and soybean<sup>3</sup>. Several specific families, including RLK-Pelle\_DLSV (158 members, the largest family in the grapevine kinome) and RLK-Pelle\_LRK10L-2 (113 members), are responsible for the large size of the RLK-Pelle group in grapevine. The RLK-Pelle\_DLSV family was also the largest family within the RLK-Pelle group of PKs in *Arabidopsis*, rice, and soybean<sup>2,3</sup>. The RLK-Pelle\_LRK10L-2 family in grapevine, however, is much larger than that in many other plant species, including *Arabidopsis*, rice, maize, and soybean, but much smaller (206) than in *Eucalyptus grandis*<sup>2</sup>.

#### The association of the grapevine kinome with plant development and plant stress response

Protein kinases have been reported to play critical regulatory roles in various biological processes in plants<sup>11,49,50</sup>. Previous studies have indicated that the grapevine CDPK gene family in the CAMK group and

MAPK in the CMGC group are involved in a variety of developmental processes and plant responses to stresses<sup>24,25</sup>. Overexpression of a grape CDPK gene in *Arabidopsis* promotes plant growth and enhances ABA sensitivity during seed germination and early seedling growth<sup>51</sup>. Transgenic *Arabidopsis* overexpressing a CDPK gene (*VaCPK20*) from *Vitis amurensis* exhibited higher tolerance to heat, salt, cold, and drought, and a similar increase in stress tolerance was observed in transgenic grape cell cultures<sup>52</sup>. The expression patterns of different grapevine kinome family genes during development and in response to a variety of stresses suggest that the role of kinase families in these processes is diverse and may occur through different mechanisms. More than half of the grapevine kinase genes in the RLK-Pelle group were downregulated in most tissues during development, similar to the situation reported in maize organs<sup>40</sup>, suggesting that they may have negative regulatory functions. Many non-RLK groups, including CAMK, CK1, CMGC, STE, and TKL, however, exhibited high expression levels in different tissue samples during development (Fig. 4), suggesting that these kinase genes may play positive regulatory roles in plant growth, as reported in soybean<sup>3</sup>.

Approximately one-third of the rice *RLK-Pelle* PK genes expressed in the embryo and endosperm were reported to be regulated by abiotic stress, with the majority of them being upregulated by salt and drought<sup>53</sup>. Overexpression of *OsSIK1*, a putative *RLK* gene from rice, improved drought and salt stress tolerance in rice<sup>54</sup>. Several *Arabidopsis* *RLK-Pelle* genes were upregulated in response to osmotic, wounding, and UV-B stresses<sup>37,55</sup>. Overexpression of an *Arabidopsis* receptor-like protein kinase 1 (RPK1) in *Arabidopsis* regulated reactive oxygen species (ROS) homeostasis and increased drought tolerance, while a mutant, *rpkl-1*, lacking RPK1 function exhibited increased water loss<sup>56</sup>. Our analysis of the *Arabidopsis* kinome families revealed different expression patterns in response to salt and drought and that a portion of the families could be significantly induced by these stresses (Supplementary Fig. S5). Among the 56 grapevine kinase families that exhibited responses to a variety of stresses (Fig. 5), 23 belong to the RLK-Pelle group. Eighteen of the 23 families exhibited either up or downregulation in response to different stress treatments, indicating that the

RLK-Pelle group may play important roles in plant responses to stresses. Interestingly, analysis of the different expression patterns under multiple treatments showed that about half of the 56 kinase families were induced by salt, and drought, which indicated that grapevine kinase genes might have important functions in responses to salt and drought stresses (Fig. 5). Based on the regulation of grapevine kinase genes by different stresses, a coexpression network of families was constructed to elucidate the relationships between different families of PK genes (Fig. 6). A kinase gene coexpression network for several kinase families in soybean<sup>3</sup> suggested that expression of those families was associated with either broad or specific stress signal transduction pathways. In our study, several grapevine RLK-Pelle families were highly coexpressed, which indicates that they may interact in plant stress response signaling. Based on the expression data of 231 grapevine PK genes under different treatments (Supplementary Table S8), the genes that were significantly induced (>two-fold,  $P < 0.05$ ) by multiple treatments would be strong candidates for further investigation.

Gene expression data can provide only indirect evidence of gene function. Thus, due to the large size of the kinase gene family in plants and the varying levels of functional redundancy within this family, confirming the functions of individual genes is a daunting challenge. Family-level analyses of expression may help to quickly identify candidate targets that are deserving of more detailed study. Since only a few grapevine kinase genes have been functionally characterized, characterization of the entire grapevine kinome, its gene expression patterns during grapevine development and in response to stress treatments, and the coexpression network of families in response to a variety of stress treatments provide fundamental information for the selection of candidate genes for further study. The 20 selected kinase genes, which were significantly induced by multiple treatments, might play key roles in stress signal transduction (Supplementary Table S9). Among the 20 genes, 7 belonged to the CAMK group and 11 were RLK-Pelle group members. These results suggest that these two groups may be involved in plant responses to a variety of stresses<sup>25,37,53</sup>. Five of the 20 genes are members of the CAMK\_CAMKL-CHK1 family, which are also known as CBL-interacting protein kinases (CIPKs) and have been demonstrated to play an essential role in plant response to a variety of stresses in other plants<sup>11,57,58</sup>. Our RNA-Seq and qRT-PCR analysis verified four PK genes (*VIT\_208s0058g01040*, *VIT\_204s0008g05770*, *VIT\_205s0020g00830*, and *VIT\_208s0058g01180*) induced by both drought and salt stresses, and three of them were

CAMK\_CAMKL-CHK1 members, suggesting that these genes might play key functions in response to environmental stresses and deserve further functional investigation (Supplementary Table S10; Fig. 7). Our information will greatly assist in developing a comprehensive understanding of this very large gene family. In addition, this study will help to clarify the roles of specific members in growth and development as well as environmental stress response and adaptation (tolerance) in grapevine.

#### Acknowledgements

This research was financially supported in part by the Priority Academic of Jiangsu Province and the Tennessee Agricultural Experiment Station Hatch Project 1009395. Kaikai Zhu was supported by the China Scholarship Council (CSC, No. 201606850052).

#### Author contributions

K.Z. and Z.-M.C. designed this research. K.Z. performed the experiments. K.Z., X. W., J.L., J.T. and Q.C. analyzed the data. K.Z. and Z.-M.C. wrote the manuscript. All authors modified the manuscript.

#### Author details

<sup>1</sup>College of Horticulture, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China. <sup>2</sup>Department of Plant Sciences, University of Tennessee, Knoxville, TN 37996, USA. <sup>3</sup>Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement, Institute of Horticulture, Jiangsu Academy of Agricultural Sciences, Nanjing, Jiangsu 210014, China. <sup>4</sup>Department of Entomology and Plant Pathology, University of Tennessee, Knoxville, TN 37996, USA. <sup>5</sup>Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

#### Conflict of interest

The authors declare that they have no conflict of interest.

**Supplementary Information** accompanies this paper at (<https://doi.org/10.1038/s41438-018-0027-0>).

Received: 8 May 2017 Revised: 13 February 2018 Accepted: 22 February 2018

Published online: 01 April 2018

#### References

- Bennett, J. Protein phosphorylation in green plant chloroplasts. *Annu. Rev. Plant Biol.* **42**, 281–311 (1991).
- Lehti-Shiu, M. D. & Shiu, S. H. Diversity, classification and function of the plant protein kinase superfamily. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **367**, 2619–2639 (2012).
- Liu, J. et al. Soybean kinome: functional classification and gene expression patterns. *J. Exp. Bot.* **66**, 1919–1934 (2015).
- Ben-David, Y., Letwin, K., Tannock, L., Bernstein, A. & Pawson, T. A mammalian protein kinase with potential for serine/threonine and tyrosine phosphorylation is related to cell cycle regulators. *EMBO J.* **10**, 317–325 (1991).
- Zulawski, M., Schulze, G., Braginets, R., Hartmann, S. & Schulze, W. X. The *Arabidopsis* kinome: phylogeny and evolutionary insights into functional diversification. *BMC Genomics* **15**, 548 (2014).
- Keates, R. Cyclic nucleotide-independent protein kinase from pea shoots. *Biochem. Biophys. Res. Commun.* **54**, 655–661 (1973).
- Lawton, M. A., Yamamoto, R. T., Hanks, S. K. & Lamb, C. J. Molecular cloning of plant transcripts encoding protein kinase homologs. *Proc. Natl Acad. Sci. USA* **86**, 3140–3144 (1989).
- Ichimura, K., Mizoguchi, T., Yoshida, R., Yuasa, T. & Shinozaki, K. Various abiotic stresses rapidly activate *Arabidopsis* MAP kinases ATMPK4 and ATMPK6. *Plant J.* **24**, 655–665 (2000).

9. Tran, L.-S. P. et al. Functional analysis of AHK1/ATHK1 and cytokinin receptor histidine kinases in response to abscisic acid, drought, and salt stress in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **104**, 20623–20628 (2007).
10. Pan, J. et al. *ZmMPK17*, a novel maize group D MAP kinase gene, is involved in multiple stress responses. *Planta* **235**, 661–676 (2012).
11. Zhu, K. et al. Evolution of an intron-poor cluster of the CIPK gene family and expression in response to drought stress in soybean. *Sci. Rep.* **6**, 28225 (2016).
12. Yang, K.-Y., Liu, Y. & Zhang, S. Activation of a mitogen-activated protein kinase pathway is involved in disease resistance in tobacco. *Proc. Natl Acad. Sci. USA* **98**, 741–746 (2001).
13. Wang, X. J. et al. Two coupled components of the mitogen-activated protein kinase cascade MdMPK1 and MdMKK1 from apple function in ABA signal transduction. *Plant Cell Physiol.* **51**, 754–766 (2010).
14. Champion, A., Kreis, M., Mockaitis, K., Picaud, A. & Henry, Y. *Arabidopsis* kinome: after the casting. *Funct. Integr. Genomics* **4**, 163–187 (2004).
15. Dardick, C., Chen, J., Richter, T., Ouyang, S. & Ronald, P. The rice kinase database. A phylogenomic database for the rice kinome. *Plant Physiol.* **143**, 579–586 (2007).
16. Chandran, A. K. N. et al. Updated Rice Kinase Database RKD 2.0: enabling transcriptome and functional analysis of rice kinase genes. *Rice* **9**, 40 (2016).
17. Caenepeel, S., Charydczak, G., Sudarsanam, S., Hunter, T. & Manning, G. The mouse kinome: discovery and comparative genomics of all mouse protein kinases. *Proc. Natl Acad. Sci. USA* **101**, 11707–11712 (2004).
18. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912–1934 (2002).
19. Hanada, K., Zou, C., Lehti-Shiu, M. D., Shinozaki, K. & Shiu, S. H. Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. *Plant Physiol.* **148**, 993–1003 (2008).
20. Hanks, S. K. & Hunter, T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576–596 (1995).
21. Fasoli, M. et al. The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. *Plant Cell* **24**, 3489–3505 (2012).
22. Ferrandino, A. & Lovisolo, C. Abiotic stress effects on grapevine (*Vitis vinifera* L.): Focus on abscisic acid-mediated consequences on secondary metabolism and berry quality. *Environ. Exp. Bot.* **103**, 138–147 (2014).
23. Jaillon, O. et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **449**, 463–467 (2007).
24. Wang, G. et al. Validation by isolation and expression analyses of the mitogen-activated protein kinase gene family in the grapevine (*Vitis vinifera* L.). *Aust. J. Grape Wine Res.* **20**, 255–262 (2014).
25. Chen, F. et al. The evolutionary history and diverse physiological roles of the grapevine calcium-dependent protein kinase gene family. *PLoS ONE* **8**, e80818 (2013).
26. Liu, J. Y., Chen, N. N., Cheng, Z. M. & Xiong, J. S. Genome-wide identification, annotation and expression profile analysis of SnRK2 gene family in grapevine. *Aust. J. Grape Wine Res.* **22**, 478–488 (2016).
27. Finn, R. D. et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285 (2016).
28. Eddy, S. R. Profile hidden Markov models. *Bioinformatics* **14**, 755–763 (1998).
29. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
30. Gasteiger, E. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784–2788 (2003).
31. Horton, P. et al. WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* **35**, W585–W587 (2007).
32. Yang, X. et al. The F-box gene family is expanded in herbaceous annual plants relative to woody perennial plants. *Plant Physiol.* **148**, 1189–1200 (2008).
33. Wang, Y., Li, J. & Paterson, A. H. MCScanX-transposed: detecting transposed gene duplications based on multiple colinearity scans. *Bioinformatics* **29**, 1458–1460 (2013).
34. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
35. Gaut, B. S., Morton, B. R., McCaig, B. C. & Clegg, M. T. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcl*. *Proc. Natl Acad. Sci. USA* **93**, 10274–10279 (1996).
36. Tang, J. et al. Characterization and co-expression analysis of WRKY orthologs involved in responses to multiple abiotic stresses in Pak-choi (*Brassica campestris* ssp. *chinensis*). *BMC Plant Biol.* **13**, 188 (2013).
37. Lehti-Shiu, M. D., Zou, C., Hanada, K. & Shiu, S. H. Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant Physiol.* **150**, 12–26 (2009).
38. Zhang, Z. et al. KaKs\_Calculator: calculating Ka and Ks through model selection and model averaging. *Genomics Proteomics Bioinformatics* **4**, 259–263 (2006).
39. Clouse, S. D. Brassinosteroid signal transduction: from receptor kinase activation to transcriptional networks regulating plant development. *Plant Cell* **23**, 1219–1230 (2011).
40. Wei, K., Wang, Y. & Xie, D. Identification and expression profile analysis of the protein kinase gene superfamily in maize development. *Mol. Breed.* **33**, 155–172 (2014).
41. Du, D., Rawat, N., Deng, Z. & Gmitter, F. G. Jr. Construction of citrus gene coexpression networks from microarray data using random matrix theory. *Hortic. Res.* **2**, 15026 (2015).
42. Liang, Y.-H. et al. Construction and validation of a gene co-expression network in grapevine (*Vitis vinifera* L.). *Hortic. Res.* **1**, 14040 (2014).
43. Shiu, S. H. et al. Comparative analysis of the receptor-like kinase family in *Arabidopsis* and rice. *Plant Cell* **16**, 1220–1234 (2004).
44. Singh, D. K. et al. The tomato kinome and the tomato kinase library ORFeome: novel resources for the study of kinases and signal transduction in tomato and solanaceae species. *Mol. Plant Microbe Interact.* **27**, 7–17 (2014).
45. Epelboin, Y. et al. The kinome of pacific oyster *Crassostrea gigas*, its expression during development and in response to environmental factors. *PLoS ONE* **11**, e0155435 (2016).
46. Wei, W., Shu, S., Zhu, W., Xiong, Y. & Peng, F. The kinome of edible and medicinal fungus *Wolfiporia cocos*. *Front. Microbiol.* **7**, 1495 (2016).
47. Kong, H. et al. Patterns of gene duplication in the plant SKP1 gene family in angiosperms: evidence for multiple mechanisms of rapid gene birth. *Plant J.* **50**, 873–885 (2007).
48. Shiu, S.-H. & Bleeker, A. B. Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in *Arab. Plant Physiol.* **132**, 530–543 (2003).
49. Becraft, P. W. Receptor kinase signaling in plant development. *Annu. Rev. Cell Dev. Biol.* **18**, 163–192 (2002).
50. Xiong, L. Disease resistance and abiotic stress tolerance in rice are inversely modulated by an abscisic acid-inducible mitogen-activated protein kinase. *Plant Cell* **15**, 745–759 (2003).
51. Yu, X. C. et al. Expression of a grape calcium-dependent protein kinase ACPK1 in *Arabidopsis thaliana* promotes plant growth and confers abscisic acid-hypersensitivity in germination, postgermination growth, and stomatal movement. *Plant Mol. Biol.* **64**, 531–538 (2007).
52. Dubrovina, A. S., Kiselev, K. V., Khristenko, V. S. & Aleynova, O. A. *VaCPK20*, a calcium-dependent protein kinase gene of wild grapevine *Vitis amurensis* Rupr., mediates cold and drought stress tolerance. *J. Plant Physiol.* **185**, 1–12 (2015).
53. Gao, L.-L. & Xue, H.-W. Global analysis of expression profiles of rice receptor-like kinase genes. *Mol. Plant* **5**, 143–153 (2012).
54. Ouyang, S. Q. et al. Receptor-like kinase *OsSK1* improves drought and salt stress tolerance in rice (*Oryza sativa*) plants. *Plant J.* **62**, 316–329 (2010).
55. Chae, L., Sudat, S., Dudoit, S., Zhu, T. & Luan, S. Diverse transcriptional programs associated with environmental stress and hormones in the *Arabidopsis* receptor-like kinase gene family. *Mol. Plant* **2**, 84–107 (2009).
56. Osakabe, Y. et al. Overproduction of the membrane-bound receptor-like protein kinase 1, RPK1, enhances abiotic stress tolerance in *Arabidopsis*. *J. Biol. Chem.* **285**, 9190–9201 (2010).
57. Zhang, H. et al. Identification and characterization of CBL and CIPK gene families in canola (*Brassica napus* L.). *BMC Plant Biol.* **14**, 8 (2014).
58. Tang, J. et al. Characterization of CIPK family in Asian pear (*Pyrus bretschneideri* Rehd) and co-expression analysis related to salt and osmotic stress responses. *Front. Plant Sci.* **7**, 1361 (2016).