PeerJ Preprints NOT PEER-REVIEWED

Genome-wide investigation of malate dehydrogenase gene family in poplar (*Populus trichocarpa*) and their expression analysis under salt stress

Xinghao Chen [1,2], **Jun Zhang** [Corresp., 1,2], **Chao Zhang** [1,2], **Shijie Wang** [1,2], **Minsheng Yang** [Corresp. 1,2]

[1] Agricultural University of Hebei, Institute of Forest Biotechnology, Forestry College, Baoding, Hebei, China

[2] Hebei Key Laboratory for Tree Genetic Resources and Forest Protection, Baoding, Hebei, China

Corresponding Authors: Jun Zhang, Minsheng Yang
Email address: zhangjunem@126.com, yangms100@126.com

Malate dehydrogenase (MDH) is widely distributed in plants and animals, and plays an important role in many metabolic processes. However, there have been few studies on *MDH* genes in poplar. In this study, 16 *MDH* gene sequences were identified from the *Populus trichocarpa* genome and renamed according to their chromosomal locations. Based on phylogenetic analysis, the *PtMDH* genes were divided into five groups, and genes that grouped together all shared the same subcellular location and had similar sequence lengths, gene structures, and conserved motifs. Two pairs of tandem duplication events and three segmental duplication events involving five genes were identified from the 15 *PtMDH* genes located on the chromosomes. Each pair of genes had a Ka/Ks ratios <1, indicating that the *MDH* gene family of *P. trichocarpa* was purified during evolution. Based on the transcriptome data of *P. trichocarpa* under salt stress and qRT-PCR verification, the expression patterns of *PtMDH* genes under salt stress were analyzed. The results showed that most of the genes were upregulated under salt stress, indicating that they play a role in the response of poplar to salt stress. The *PtmMDH1* gene can be used as an important salt-tolerant candidate gene for further investigations of molecular mechanisms. This study lays the foundation for functional analysis of *MDH* genes and genetic improvement in poplar.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27742v1 | CC BY 4.0 Open Access | rec: 18 May 2019, publ: 18 May 2019

PeerJ Preprints NOT PEER-REVIEWED

1 # Genome-wide investigation of malate dehydrogenase
2 # gene family in poplar (*Populus trichocarpa*) and their
3 # expression analysis under salt stress

4

5 Xinghao Chen[1,2], Jun Zhang[1,2], Chao Zhang[1,2] , Shijie Wang[1,2] and Minsheng Yang[1,2]

6

7 [1] Institute of Forest Biotechnology, Forestry College, Agricultural University of Hebei, Baoding,
8 Hebei, China

9 [2] Hebei Key Laboratory for Tree Genetic Resources and Forest Protection, Baoding, Hebei,
10 China

11

12 Corresponding Author:

13 Jun Zhang

14 No. 289, Lingyusi Street, Baoding, Hebei, 071000, China

15 zhangjunem@126.com

16

17 Minsheng Yang

18 No. 289, Lingyusi Street, Baoding, Hebei, 071000, China

19 yangms100@126.com

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27742v1 | CC BY 4.0 Open Access | rec: 18 May 2019, publ: 18 May 2019

PeerJ Preprints NOT PEER-REVIEWED

20 # Genome-wide investigation of malate dehydrogenase
21 gene family in poplar (*Populus trichocarpa*) and their
22 expression analysis under salt stress

23

24 Xinghao Chen[1,2], Jun Zhang[1,2], Chao Zhang[1,2] , Shijie Wang[1,2] and Minsheng Yang[1,2]

25

26 [1] Institute of Forest Biotechnology, Forestry College, Agricultural University of Hebei, Baoding,
27 Hebei, China

28 [2] Hebei Key Laboratory for Tree Genetic Resources and Forest Protection, Baoding, Hebei,
29 China

30

31 Corresponding Author:

32 Jun Zhang

33 No. 289, Lingyusi Street, Baoding, Hebei, 071000, China

34 zhangjunem@126.com

35

36 Minsheng Yang

37 No. 289, Lingyusi Street, Baoding, Hebei, 071000, China

38 yangms100@126.com

39

40 ## Abstract

41 Malate dehydrogenase (MDH) is widely distributed in plants and animals, and plays an
42 important role in many metabolic processes. However, there have been few studies on *MDH*
43 genes in poplar. In this study, 16 *MDH* gene sequences were identified from the *Populus*
44 *trichocarpa* genome and renamed according to their chromosomal locations. Based on
45 phylogenetic analysis, the *PtMDH* genes were divided into five groups, and genes that grouped
46 together all shared the same subcellular location and had similar sequence lengths, gene
47 structures, and conserved motifs. Two pairs of tandem duplication events and three segmental

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27742v1 | CC BY 4.0 Open Access | rec: 18 May 2019, publ: 18 May 2019

48 duplication events involving five genes were identified from the 15 *PtMDH* genes located on the
49 chromosomes. Each pair of genes had a Ka/Ks ratios <1, indicating that the *MDH* gene family of
50 *P. trichocarpa* was purified during evolution. Based on the transcriptome data of *P. trichocarpa*
51 under salt stress and qRT-PCR verification, the expression patterns of *PtMDH* genes under salt
52 stress were analyzed. The results showed that most of the genes were upregulated under salt
53 stress, indicating that they play a role in the response of poplar to salt stress. The *PtmMDH1*
54 gene can be used as an important salt-tolerant candidate gene for further investigations of
55 molecular mechanisms. This study lays the foundation for functional analysis of *MDH* genes and
56 genetic improvement in poplar.

57

58 Introduction

59 Malate dehydrogenase (MDH, EC1.1.1.37), which is widely distributed in animals and plants,
60 catalyzes the interconversion of malate and oxaloacetate (OAA) and plays important roles in
61 several metabolic pathways including the tricarboxylic acid cycle, glyoxylate bypass,
62 photosynthesis, and the C4-dicarboxylic acid cycle (Selinski et al., 2014; Gietl, 1992). MDH is a
63 multimeric enzyme that usually occurs as dimers or tetramers consisting of the same or similar
64 subunits; an exception is the MDH from *Nitzschia alba*, which is an octamer (Yueh, Chung &
65 Lai, 1989). Each subunit of MDH contains functional and structural domains: the nicotinamide
66 adenine dinucleotide (NAD)-binding domain occupies the N-terminus, and the C-terminal
67 domain is the substrate-binding site; the necessary amino acid residues for catalysis, as well as
68 the active site of the enzyme, exist in the spaces between these two domains (Musrati et al.,
69 1998). Higher plants contain multiple forms of MDH, which differ in coenzyme specificity,
70 subcellular localization, and physiological functions (Gietl, 1992). According to the coenzyme
71 specificity and subcellular location, MDH is divided into NAD-dependent MDH (NAD-MDH)
72 and NADP-dependent MDH (NADP-MDH). NAD-MDH, with the coenzyme NAD+, is
73 distributed in mitochondria, cytoplasm, microbodies (peroxisomes and glyoxysomes), and
74 chloroplasts, and the molecular weight of each subunit ranges from 32–37 kDa. NADP-MDH,
75 with the coenzyme NADP+, is distributed in chloroplasts, and the molecular weight of each
76 subunit ranges from 42–43 kDa (Ding & Ma, 2004; Tomaz et al., 2010).

77 Many studies have shown that MDHs in plants play crucial roles in seed germination, plant
78 growth, pollen and fruit development, and the response to abiotic stress (Beeler et al., 2014;
79 Rudrappa, 2008). The expression of the cytosolic NAD-dependent *MDH* gene in apple is
80 positively correlated with growth viability and metabolic activity, and plays an important role in
81 plant growth and response to low temperatures and salt stress (Yao et al., 2011). Overexpression
82 of the NADP-dependent *MDH* gene in chloroplasts significantly affects the redox state of
83 ferredoxin in transgenic potato plants (Backhausen et al., 1998). The activity of the
84 mitochondrial NAD-dependent *MDH* gene in strawberry increases significantly with the maturity
85 of the fruit, indicating that the gene is closely related to the growth and development of the fruit

(Iannetta, et al., 2004). The NAD-dependent *MDH* gene isolated from cucumber was shown to play roles in the peroxisome and glyoxysome, and its expression increased significantly during seed germination (Kim & Smith, 1994). At present, studies on MDH are becoming more and more extensive, but most focus on crops and horticultural plants, and studies that focus on poplar are relatively rare.

The poplar has a long history of cultivation and is widely used as an important industrial material as well as in ecological protection as an economical forest tree species (Tun et al., 2018). With the increase of soil salinization, salt stress has become an important factor restricting the growth of forest trees (Zhou et al., 2018); therefore, research on the functions of genes related to salt tolerance has received increasing attention. In this study, we applied bioinformatics methods to identify *MDH* gene family members from the genome of *Populus trichocarpa*, and analyzed the phylogeny, gene structure, chromosomal location, and duplication of each member. Using quantitative real-time reverse-transcription PCR (qRT-PCR) and transcriptome data, the expression patterns of *MDH* gene family members under salt stress were analyzed. The results lay the foundation for further analysis of the function of *MDH* genes in poplar.

Materials & Methods

Identification of MDH family members in the *P. trichocarpa* genome

The whole genome and protein sequences of *P. trichocarpa* were downloaded from the Ensembl Plants database (http://plants.ensembl.org/index.html). First, the identified MDH protein sequences from *Arabidopsis thaliana* (downloaded from the Arabidopsis Information Resource (TAIR); https://www.arabidopsis.org/) were used as queries in BLASTP searches against the *P. trichocarpa* genome databases with an e-value cutoff of 1.0 (Imran, Tang & Liu, 2016). The redundant sequences were manually removed, and all candidate genes were analyzed using InterProScan (http://www.ebi.ac.uk/interpro/search/sequence-search) and SMART (http://smart.emblheidelberg.de/) to confirm the presence of the MDH domain (PF00056, PF02866, and PS00068) (Quevillon et al., 2005; Letunic & Bork, 2018). Finally, the sequence lengths, isoelectric points, and molecular weights of the MDH protein sequences from *P. trichocarpa* were calculated using ExPASy (http://cn.expasy.org/tools), and the online software WoLF PSORT (http://www.genscript.com/wolf-psort.html) and EuLoc (http://euloc.mbc.nctu.edu.tw/) were used to predict the subcellular localization of MDH proteins in *P. trichocarpa*.

Phylogenetic analysis and classification of the *MDH* gene family in *P. trichocarpa*

The MDH protein sequences from *P. trichocarpa* and *A. thaliana* were used for phylogenetic analysis. First, a multiple sequence alignment of MDH protein sequences was generated using

122 ClustalW in MEGA 7 with the default parameters. A neighbor-joining phylogenetic tree was
123 constructed based on the alignment results with the Poisson model, pairwise deletion, and 1,000
124 bootstrap replications (Xie et al., 2018). Finally, the MDH proteins were divided into groups
125 according to the topology of the phylogenetic tree.

126 **Structure and motif analysis of the *MDH* gene family in *P. trichocarpa***

127 Structural information on the *MDH* gene family members was extracted from the genome
128 annotation file of *P. trichocarpa*. The online Gene Structure Display Server (GSDS;
129 http://gsds.cbi.pku.edu.cn) was used to map the structures of the *PtMDH* genes (Guo, Zhu &
130 Chen, 2007). MEME software was used to analyze the motifs of PtMDH protein sequences, and
131 the parameters were set as follows: amino acid length, 6–50; number of repeats of the motif,
132 arbitrary; threshold number of motif discovery, 10 (Bailey et al., 2006).

133 **Chromosomal localization and duplication analysis of *MDH* genes in *P. trichocarpa***

134 Based on the genome annotation file of *P. trichocarpa*, positional information on *PtMDH* genes
135 was obtained, and the chromosomal localizations of *PtMDH* genes were mapped using
136 MapChart software. *PtMDH* gene duplication events were analyzed using the Multiple
137 Collinearity Scan toolkit (MCScanX). The definition criteria for segmental gene duplication
138 were as follows: 1) the length of the shorter aligned sequence covered >70% of the longer
139 sequence, and 2) the similarity of the two aligned sequences was >70% (Wang et al., 2012; Gu et
140 al., 2002). In addition, homologous genes located on the same chromosome with no more than
141 one gene between them were identified as tandem duplication genes (Wu et al., 2014). Circos
142 was used to map the results of *MDH* gene duplication events as images (Krzywinski et al., 2009).
143 KaKs_Calculator 1.2 was used to calculate non-synonymous (Ka) and synonymous (Ks)
144 substitutions of each duplicated MDH gene (Zhang et al., 2006).

145 **Transcriptomic analysis of *PtMDH genes* under salt stress**

146 The expression levels of *MDH* gene family members were extracted from the transcriptome data
147 of *Populus* × *euramericana* cv. '74/76' treated with different salt concentrations (0, 3‰, and 6‰)
148 obtained by high-throughput sequencing in a previous study (Chen et al., 2018). After the data
149 were $\log_2^{(1+RPKM)}$ transformed, a gene expression heat map was generating using MEV 4.0 (Saeed
150 et al., 2003). According to differences in gene expression under salt stress, the 16 *PtMDH* genes
151 were classified and their expression patterns were explored.

152 **Plant material and salt stress treatment**

153 Tissue culture seedlings of *Populus* × *euramericana* cv. '74/76' were used as experimental
154 material. Stem segments with two leaves were inoculated into 1/2 MS medium containing 0
155 mmol/L or 100 mmol/L NaCl (1.5% sucrose + 0.65% agar + 0.3 mg/L IBA). The tissues were
156 cultured at 25°C under a light intensity of 2,000 lux with a 14-h light/10-h dark photoperiod.

157 Root, stem, and leaf tissues were collected at 15, 20, 25, and 30 days, and immediately frozen in
158 liquid nitrogen and stored at −80°C for subsequent analysis.

159 **RNA extraction and qRT-PCR analysis**

160 The EASYspin plant RNA rapid extraction kit was used to extract RNA from tissue samples
161 collected at different time points. The concentration and quality of RNA were determined using a
162 NanoDrop One Microvolume UV-Vis spectrophotometer and an Agilent 2100 Bioanalyzer.
163 cDNA was synthesized via reverse transcription using the First Strand cDNA Synthesis Kit with
164 the ReverTra Ace qPCR RT Master Mix. Primer Premier 5.0 was used to design primers for the
165 *PtMDH* genes (Table 1). qRT-PCR was performed on an Agilent Mx3005P using SYBR Green
166 chemistry. The *PtRG5* gene encoding a poplar ribosomal L29e family protein was used as an
167 internal control. The reaction was carried out as follows: 95°C for 30s, followed by 50 cycles of
168 95°C for 10 s, 56°C for 30 s, and 72°C for 34 s. Each reaction was performed with three
169 biological replicates and the expression levels of the genes were calculated using the $2^{-\Delta\Delta Ct}$
170 method (Livak & Schmittgen, 2001).

171

172 Results

173 **Identification and analysis of *MDH* genes in *P. trichocarpa***

174 After removal of redundant sequences and database validation, a total of 16 *MDH* genes were
175 identified in the genome of *P. trichocarpa*. The 16 *PtMDH* genes were renamed according to
176 their chromosomal positions and the results of subcellular localization (Table 2). Except for
177 *PtmMDH5*, the remaining 15 *PtMDH* genes were mapped to chromosomes. The lengths of the
178 *PtMDH* genes varied; *PtmMDH4* was the shortest at 1,491 bp, whereas *PtchMDH2* was the
179 longest at 4,762 bp. In contrast, the lengths of the coding sequences and protein sequences were
180 less varied, with averages of 1,099 bp and 365 amino acids, respectively. The molecular weights
181 of the PtMDH proteins ranged from 27.68 KDa to 49.76 KDa, with an average of 38.87 KDa.
182 The isoelectric points ranged from 5.87 to 8.93, with an average of 7.96. Of the 16 *PtMDH*
183 genes, five were localized to the mitochondrion, three to the peroxisome, four to the cytoplasm,
184 and four to the chloroplast.

185 **Phylogenetic analysis and classification of *MDH* genes in *P. trichocarpa***

186 To identify similar regions and conserved sites among *PtMDH* genes and to clarify their
187 evolutionary relationships with different species, we performed multiple sequence alignment
188 with the 16 PtMDH protein sequences and nine AtMDH protein sequences, and built a
189 phylogenetic tree based on the alignment results. Based on the analysis, the 25 MDH sequences
190 were divided into five groups, and the *PtMDH* genes in each group shared the same subcellular
191 location, indicating that the subcellular localization of MDH proteins is closely related to

192 information contained in the sequence. It is worth noting that although Group I and Group II
193 share the same subcellular location, they did not cluster together. According to the National
194 Center for Biotechnology Information database, AtchMDH1 (in Group I) is a NAD-dependent
195 MDH, whereas AtchMDH2 (in Group II) is a NADP-dependent MDH, indicating that among the
196 four *PtchMDH* genes, *PtchMDH1*, *PtchMDH3*, and *PtchMDH4*, whose coenzyme is NAD+, are
197 closely related to *MDH* genes that localize to the mitochondrion and peroxisome, whereas
198 *PtchMDH2*, whose coenzyme is NADP+, is closely related to *MDH* genes that localize to the
199 cytoplasm (Fig. 1).

200 **Gene structure and motif analysis of MDH members in *P. trichocarpa***

201 We analyzed the structures of the *PtMDH* genes to further understand their structural evolution.
202 The number of introns in the *PtMDH2* genes ranged from 0 to 12, with the greatest number of
203 introns in the NADP-dependent *PtchMDH2* gene, and no introns in the NAD-dependent
204 *PtchMDH1*, *PtchMDH3*, and *PtchMDH4* genes. Although all four of their proteins localized to
205 the chloroplast, the number of introns in the genes varied greatly, and may be related to their
206 different coenzymes and functions. *PtMDH* genes in the same group had similar numbers of
207 exons and introns, whereas the difference between groups was significant (Fig. 2).

208 Ten conserved motifs were found in the PtMDH proteins by MEME analysis. The conserved
209 motifs ranged from 21 to 100 amino acids in length. Each *PtMDH* gene contained 4–5 conserved
210 motifs, and all 16 contained motif 1. Based on analysis using the Pfam and SMART databases,
211 motif 1 corresponds to the C-terminal domain. The NAD-binding domains of the PtMDH protein
212 sequences consist entirely of motifs 2 and 4 or motif 8. In the same group, the types and
213 quantities of motifs were generally the same. Some conserved motifs were exclusive to a certain
214 group, such as motif 8 (Group I) and motif 5 (Group III). The motifs are structural components
215 of proteins with specific spatial conformations and functions, indicating differences in evolution
216 and function between MDH members with different motifs.

217 **Chromosomal localization and duplication analysis of *MDH* genes in *P. trichocarpa***

218 Figure 3 shows that except for the *PtmMDH5* gene, which was not located on a chromosome, the
219 remaining 15 *PtMDH* genes are unevenly distributed on nine chromosomes of *P. trichocarpa*.
220 Chromosome 17 contained the most *PtMDH* genes (three), whereas chromosomes 7, 9, 10, and
221 11 each contain only one.

222 As tandem duplication and segmental duplication affect the formation of the gene family to a
223 certain extent and are the driving force of genome evolution, we analyzed the duplication events
224 of *PtMDH* genes. Four *PtMDH* genes, *PtcyMDH1/PtcyMDH2* and *PtchMDH3/ PtchMDH4*,
225 were identified as tandem duplication genes on chromosomes 2 and 17, respectively. In addition,
226 three segmental duplication events involving five *MDH* genes were also identified (Fig. 4). To
227 better understand the evolutionary pressure on the *PtMDH* genes, we calculated the Ka/Ks ratios

PeerJ Preprints NOT PEER-REVIEWED

228 of the duplicated gene pairs. All duplicated gene pairs had Ka/Ks ratios <1, indicating that the
229 *MDH* gene family of *P. trichocarpa* has experienced purifying selection.

230 **Expression pattern analysis of the *MDH* gene family of *P. trichocarpa* under salt stress**

231 Based on transcriptome data from *Populus* × *euramericana* cv. '74/76' treated with different salt
232 concentrations, we analyzed the expression changes of *P. trichocarpa MDH* genes. Fourteen
233 genes showed significant changes in their expression levels, indicating that these genes are
234 involved in the response of poplar to salt stress. The expression patterns of the *PtMDH* gene
235 family members were roughly divided into five categories: class I showed decreased gene
236 expression in response to salt stress; class II showed increased expression levels under low salt
237 stress, and decreased expression under high salt stress; class III showed decreased expression
238 under low salt stress, and increased expression under high salt stress; class IV showed no
239 changes in expression under low salt stress, but increased expression under high salt stress; and
240 class V genes were not expressed under salt stress, and may be pseudogenes that formed during
241 the evolutionary process of the *MDH* gene family but lost function due to sequence changes, or
242 they may exhibit spatiotemporal expression specificity (Fig. 5).

243 To examine temporal and tissue-specific expression, we analyzed the expression levels of six
244 *PtMDH* genes in the roots, stems and, leaves of class IV genes under salt stress by qRT-PCR.
245 Most of the genes were significantly upregulated in various tissues on the 25th and 30th day after
246 salt stress. At the four time points, in various tissues, the expression levels of *PtmMDH1* were
247 significantly higher than those in the control, indicating that salt treatment had a significant
248 induction effect on the *PtmMDH1* gene. Therefore, the *PtmMDH1* gene can be used as an
249 important salt tolerant candidate gene for functional verification in subsequent experiments (Fig.
250 6).

251

252 ## Discussion

253 MDH is a highly active enzyme in plants and plays an indispensable role in many metabolic
254 processes (Yao et al., 2011). At present, most studies on *MDH* genes are focused in crops and
255 horticultural plants, such as cotton (Imran, Tang & Liu, 2016; Imran & Liu, 2016), corn (Metzler
256 et al., 1989), wheat (Ding & Ma, 2004), apple (Yao et al., 2011), grape (Sweetman et al., 2009),
257 etc. By contrast, there is very little research on the function of *MDH* genes in poplar, a model
258 plant for forest genomics. We identified 16 *MDH* gene sequences in the genome of *P.*
259 *trichocarpa* using bioinformatics methods, which is more than the number identified in
260 *Gossypium arboreum* (Imran & Liu, 2016), but less than in *Gossypium hirsutum* (Imran, Tang &
261 Liu, 2016). Based on phylogenetic analysis with *A. thaliana*, the 16 *PtMDH* genes were divided
262 into five groups, with the genes in each group sharing the same subcellular location. The four
263 *PtchMDH* genes, which localize to the chloroplast, were divided into two groups, likely because
264 they have different coenzymes, resulting in differences in function.

265 A total of 10 motifs were identified in the *MDH* gene family of *P. trichocarpa*. Although the
266 sequences in different groups have different types and numbers of motifs, they all have motifs
267 corresponding to the NAD-binding domain and C-terminal domain of MDH proteins. For
268 example, in groups I, IV, and V, motifs 2 and 4 correspond to the NAD-binding domain, whereas
269 motifs 1 and 3 correspond to the C-terminal domain; in group III, motif 6 corresponds to the
270 NAD-binding domain, whereas motifs 1 and 5 correspond to the C-terminal domain. Motif 1 is
271 present in every MDH sequence because it contains an active site that is an indispensable
272 component of MDH sequences and plays a catalytic role when the substrate binds to MDH. A
273 previous study showed that the critical residue of the active site distinguishes between malate
274 dehydrogenase and lactate dehydrogenase, two enzymes that share high sequence similarity
275 (Hannenhalli & Russell, 2000). The five groups of *PtMDH* genes have different intron-exon
276 structures and different numbers of introns. Unlike the other groups, the genes in group I have no
277 introns, consistent with the results of a study on the cotton *MDH* gene family by Imran et al.
278 (Imran, Tang & Liu, 2016). Ren et al. reported that highly expressed genes have more and longer
279 introns than genes expressed at low levels, suggesting that the *PtMDH* genes in group I are
280 expressed at relatively low levels in response to biotic or abiotic stresses (Ren et al., 2006).

281 Gene duplication events, mainly including tandem duplication, segmental duplication, and
282 transposition events, are important ways to expand gene families (Kong et al., 2007). Among the
283 15 *PtMDH* genes distributed unevenly on nine chromosomes, two pairs of tandem duplication
284 events and three segmental duplication events involving five genes were identified. Based on
285 analysis of transcriptome data, the expression patterns of different *PtMDH* genes under salt stress
286 vary, indicating that the functions of some *PtMDH* genes may have changed during the evolution
287 of the gene family. The *PtcyMDH2* gene may have lost its function and become a pseudogene
288 after being formed from the *PtcyMDH1* gene, which often occurs in multigene families in
289 eukaryotes (Maestre et al., 1995).

290 According to the expression levels of *PtMDH* genes in transcriptome data, the genes were
291 divided into five categories, each of which exhibits a certain expression pattern. The expression
292 levels of most *PtMDH* genes increased significantly in response to salt stress, but at different salt
293 concentrations, indicating the existence of multiple molecular mechanisms through which
294 *PtMDH* genes respond to salt stress. qRT-PCR was used to detect the transcription levels of
295 several *PtMDH* genes in various tissues under salt stress. The expression of most *PtMDH* genes
296 increased significantly compared to the control during the late stage of treatment, especially the
297 *PtmMDH1* gene, the expression of which was higher than that of the control at multiple time
298 points in multiple tissues, indicating that this gene plays an important role in the response of
299 poplar to salt stress, and serves as an important salt tolerance candidate gene for further research
300 on molecular mechanisms.

301

302 **Conclusions**

303 A total of 16 *MDH* gene sequences were identified in the *P. trichocarpa* genome. Based on the
304 phylogenetic tree constructed with *A. thaliana* sequences, the *16 PtMDH* genes were divided into
305 five groups; genes in the same group shared the same subcellular location and had similar gene
306 structures. Analysis of gene duplication events and evolutionary selection revealed the formation
307 and evolutionary processes of the *MDH* gene family in *P. trichocarpa*. The expression patterns
308 and changes in *PtMDH* genes under salt stress were further analyzed based on published
309 transcriptome data and qRT-PCR. The *PtmMDH1* gene can serve as a candidate gene for
310 studying salt tolerance mechanisms. This study lays the foundation for further analysis of the
311 function of *MDH* genes in poplar.

312

313

314

315 ## References

316 Backhausen JE, Emmerlich A, Holtgrefe S, Horton P, Nast G, Rogers JJM, Muller-Rober B,
317 Scheibe R. 1998. Transgenic potato plants with altered expression levels of chloroplast
318 NADP-malate dehydrogenase: interactions between photosynthetic electron transport and
319 malate metabolism in leaves and in isolated intact chloroplasts. *Planta* 207:105-114 DOI
320 10.2307/23385418.

321 Bailey TL, Williams N, Misleh C, Li WW. 2006. MEME: discovering an analyzing DNA and
322 protein sequence motifs. *Nucleic Acids Research* 34:369-373 DOI 10.1093/nar/gkl198.

323 Beeler S, Liu HC, Stadler M, Schreier T, Eicke S, Lue WL, Truernit E, Zeeman SC, Chen J,
324 Kotting O. 2014. Plastidial NAD-dependent malate dehydrogenase is critical for embryo
325 development and heterotrophic metabolism in Arabidopsis. *Plant Physiology* 164((3):1175-
326 1190 DOI 10.1104/pp.113.233866.

327 Chen P, Zuo L, Yu X, Dong Y, Zhang S, Yang M. 2018. Response mechanism
328 in *Populus×euramericana* cv.'74/76' revealed by RNA-seq under salt stress. *Acta*
329 *Physiologiae Plantarum* 40:96 DOI 10.1007/s11738-018-2676-x.

330 Ding Y, Ma QH. 2004. Characterization of cytosolic malate dehydrogenase cDNA which
331 encodes an isozyme toward oxaloacetate reduction in wheat. *Biochimie* 86:509-518 DOI
332 10.1016/j.biochi.2004.07.011.

333 Gietl C. 1992. Malate dehydrogenase isoenzymes: Cellular locations and role in the flow of
334 metabolites between the cytoplasm and cell organelles. *Biochimica et Biophysica Acta (BBA)*
335 *- Bioenergetics* 1100(3):217-234 DOI 10.1016/0167-4838(92)90476-T.

336 Gu Z, Cavalcanti A, Chen FC, Bouman P, Li WH. 2002. Extent of gene duplication in the
337 genomes of Drosophila, nematode, and yeast. *Molecular Biology and Evolution* 19(3):256-
338 262 DOI 10.1093/oxfordjournals.molbev.a004079

339 Guo AY, Zhu QH, Chen X, Luo JC. 2007. GSDS: a gene structure display server. *Hereditas*
340 *(Beijing)* 29(8):1023-1026

341 Hannenhalli SS, Russell RB. 2000. Analysis and prediction of functional sub-types from protein
342 sequence alignments. *Journal of Molecular Biology* 303(1):61-76 DOI
343 10.1016/j.etap.2004.12.050.

344 Iannetta PPM, Escobar NM, Ross HA, Souleyre EJ, Hancock RD, Witte CP, Davies HV. 2004.
345 Identification, cloning and expression analysis of strawberry (Fragaria × ananassa)
346 mitochondrial citrate synthase and mitochondrial malate dehydrogenase. *Plant Physiology*
347 121(1):15-26 DOI 10.1111/j.0031-9317.2004.00302.x.

348 Imran M, Liu JY. 2016. Genome-wide identification and expression analysis of the malate
349 dehydrogenase gene family in *Gossypium arboreum*. *Pakistan Journal of Botany*
350 48(3):1081-1090

351 Imran M, Tang K, Liu JY. 2016. Comparative genome-wide analysis of the malate dehydrogense
352 gene families in cotton. *PLOS ONE* 11(11):e0166341 DOI 10.1371/journal.pone.0166341.

353 Kim DJ, Smith SM. 1994. Expression of a single gene encoding microbody NAD-malate
354 dehydrogenase during glyoxysome and peroxisome development in cucumber. *Plant*
355 *Molecular Biology* 26:1833-1841 DOI 10.1007/BF00019496.

356 Kong H, Landherr LL, Frohlich MW, Leebens-Mack J, Ma H, DePamphilis CW. 2007. Patterns
357 of gene duplication in the plant SKP1 gene family in angiosperms: evidence for multiple
358 mechanisms of rapid gene birth. *Plant Journal* 50(5):873–885 DOI 10.1111/j.1365-
359 313X.2007.03097.x.

360 Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA.
361 2009. Circos: An information aesthetic for comparative genomics. *Genome Research*
362 19:1639-1645 DOI 10.1101/gr.092759.109.

363 Letunic I, Bork P. 2018. 20 years of the SMART protein domain annotation resource. *Nucleic*
364 *Acids Research* 46:493-496 DOI 10.1093/nar/gkx922.

365 Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time
366 quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25:402-408 DOI
367 10.1006/meth.2011.1262.

368 Maestre J, Tchenio T, Dhellin O, Heidmann T. 1995. mRNA retroposition in human cells:
369 processed pseudogene formation. *The Embo Journal* 14(24):6333-6338 DOI
370 10.1002/j.1460-2075.1995.tb00324.x.

371 Metzler MC, Rothermel BA, Nelson T. 1989. Maize NADP-malate dehydrogenase: cDNA
372 cloning, sequence, and mRNA characterization. *Plant Molecular Biology* 12(6):713-722
373 DOI 10.1007/BF00044162.

374 Musrati RA, Kollarova M, Mernik N, Mikulasova D. 1998. Malate dehydrogenase: distribution,
375 function and properties. *General Physiology and Biophysics* 17(3):193-210 DOI
376 10.1023/A:1007730308184.

377 Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. 2005.
378 InterProScan: protein domains identifier. *Nucleic Acids Research* 33:116-120 DOI
379 10.1093/nar/gki442.

380 Ren XY, Vorst O, Fiers MWEJ, Stiekema WJ, Nap JP. 2006. In plants, highly expressed genes
381 are the least compact. *Trends in Genetics* 22:528–532 DOI 10.1016/j.tig.2006.08.008.

382 Rudrappa T, Czymmek KJ, Pare PW, Bais HP. 2008. Root-secreted malic acid recruits beneficial
383 soil bacteria. *Plant Physiology* 148(3):1547-1556 DOI 10.1104/pp.108.127613.

384 Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T,
385 Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E,
386 Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J. 2003. TM4: a free, open-source
387 system for microarray data management and analysis. *Biotechniques* 34(2):374-378 DOI
388 10.2144/03342mt01.

389 Selinski J, König N, Wellmeyer B, Hanke GT, Linke V, Neuhaus HE, Scheibe R. 2014. The
390 plastid-localized NAD-dependent malate dehydrogenase is crucial for energy homeostasis
391 in developing Arabidopsis thaliana seeds. *Molecular Plant* 7(1):170-186 DOI
392 10.1093/mp/sst151.

393 Sweetman C, Deluc LG, Cramer GR, Ford CM, Soole KL. 2009. Regulation of malate
394 metabolism in grape berry and other developing fruits. *Phytochemistry* 70(11-12):1329-
395 1344 DOI 10.1016/j.phytochem.2009.08.006.

396 Tomaz T, Bagard M, Pracharoenwattana I, Lindén P, Lee CP, Carroll AJ, Stroher E, Smith SM,
397 Gardestrom P, Millar AH. 2010. Mitochondrial malate dehydrogenase lowers leaf
398 respiration and alters photorespiration and plant growth in Arabidopsis. *Plant Physiology*
399 154(3):1143-1157 DOI 10.2307/25758664.

400 Tun TN, Guo J, Fang SZ, Tian Y. 2018. Planting spacing affects canopy structure, biomass
401 production and stem roundness in poplar plantations. *Scandinavian Journal Forest*
402 *Research* 33(5):464-474 DOI 10.1080/02827581.2018.1457711.

403 Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee T, Jin H, Marler B, Guo H, Kissinger
404 JC, Paterson AH. 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene
405 synteny and collinearity. *Nucleic Acids Research* 40(7):e49 DOI 10.1093/nar/gkr1293.

406 Wu N, Zhu Y, Song W, Li Y, Yan Y, Hu Y. 2014. Unusual tandem expansion and positive
407 selection in subgroups of the plant GRAS transcription factor superfamily. *BMC Plant*
408 *Biology* 14:373-393 DOI 10.1186/s12870-014-0373-5.

409 Xie T, Chen C, Li C, Liu J, Liu C, He Y. 2018. Genome-wide investigation of WRKY gene
410 family in pineapple: evolution and expression profiles during development and stress. *BMC*
411 *Genomics* 19(1):490-507 DOI 10.1186/s12864-018-4880-x.

412 Yao YX, Dong QL, Zhai H, You CX, Hao YJ. 2011. The functions of an apple cytosolic malate
413 dehydrogenase gene in growth and tolerance to cold and salt stresses. *Plant Physiology and*
414 *Biochemistry* 49(3):257-264 DOI 10.1016/j.plaphy.2010.12.009.

415 Yueh AY, Chung CS, Lai YK. 1989. Purification and molecular properties of malate
416 dehydrogenase from the marine diatom Nitzschia alba. *Biochemical Journal* 258(1):221-
417 228 DOI 10.1016/0006-291X(89)92830-1.

418 Zhang Z, Li J, Zhao XQ, Wang J, Wong GKS, Yu J. 2006. KaKs_Calculator: calculating Ka and
419 Ks through model selection and model averaging. *Genomics, Proteomics & Bioinformatics*
420 4(4):259-263 DOI 10.1016/S1672-0229(07)60007-2.

421 Zhou Y, Tang N, Huang L, Zhao Y, Tang X, Wang K. 2018. Effects of salt stress on plant
422 growth, antioxidant capacity, glandular trichome density, and volatile exudates of
423 *schizonepeta tenuifolia* Briq.. *International Journal Molecular Sciences* 19(1):252-266 DOI
424 10.3390/ijms19010252.

Table 1 (on next page)

Primer sequences for quantitative real-time reverse-transcription PCR (qRT-PCR).

PeerJ Preprints NOT PEER-REVIEWED

Gene name	Forward primer	Reverse primer
PtpMDH2	CTTCTGCTGTGGTTCGTGGGTTCTT	TCATCCCTCGTCATTCCTGGTTTTC
PtcyMDH3	GACAAGACTGGACCATAACAGGGCACT	TTGACATCAGGGTACTGGGAAGACGA
PtcyMDH4	GAATGGTGTGAAGATGGAGTTGGTA	GAGCAGAAGCCTGTGACTTGTAAAT
PtmMDH1	CAAGATGGTGGGACAGAAGTTGTGGAA	TCACCCTGGATGCGAAGAAAGGTAGTT
PtchMDH1	AAGCCAAAAGTCACAGTATCACCTAAACC	GCTGAAACCAATGGAGACATCTTAACTAGA
PtchMDH4	CTAAACCCACAGGCATCCTACAAAG	ACTTGCGAGGGAGTGTTACAGTGAC
PtRG5	CCCAGAGCCGCACCAACT	TGGGTTTCTTGATGCCATTTTG

1 Table 1. Primer sequences for quantitative real-time reverse-transcription PCR (qRT-PCR).

2

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27742v1 | CC BY 4.0 Open Access | rec: 18 May 2019, publ: 18 May 2019

PeerJ Preprints NOT PEER-REVIEWED

Table 2 (on next page)

Characteristics of the malate dehydrogenase (*MDH*) gene family in *Populus trichocarpa*.

Chr: chromosome number; Un: unknown; CDS: coding sequence length (nucleotides); aa: amino acid length; MW: molecular weight; pI: isoelectric point.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27742v1 | CC BY 4.0 Open Access | rec: 18 May 2019, publ: 18 May 2019

Name	Gene ID	Chr	Genomic location	CDS	Proteins			Subcellular location
					aa	MW	pI	
PtpMDH1	POPTR_001G287400	1	29339914:29342693	1071	356	37.62	8.41	Peroxisome
PtpMDH2	POPTR_007G009100	7	697553:701897	1065	354	37.51	8.4	Peroxisome
PtpMDH3	POPTR_009G081600	9	7755093:7758631	1065	354	37.49	8.61	Peroxisome
PtcyMDH1	POPTR_002G141700	2	10519570:10522106	1095	364	39.56	6.02	Cytoplasm
PtcyMDH2	POPTR_002G141900	2	10525190:10527138	1005	334	36.58	5.87	Cytoplasm
PtcyMDH3	POPTR_008G166800	8	11393003:11395488	1179	392	42.9	7.07	Cytoplasm
PtcyMDH4	POPTR_010G071000	10	9779104:9782553	999	332	35.66	6.19	Cytoplasm
PtmMDH1	POPTR_001G376500	1	39229408:39233759	1035	344	36.01	8.75	Mitochondria
PtmMDH2	POPTR_004G054200	4	4248885:4250949	1047	348	36.34	8.12	Mitochondria
PtmMDH3	POPTR_011G096300	11	11739063:11743279	1104	367	38.42	8.82	Mitochondria
PtmMDH4	POPTR_017G152000	17	15815003:15816493	789	262	27.68	9.81	Mitochondria
PtmMDH5	POPTR_T143500	Un	4048:6131	1047	348	36.47	8.69	Mitochondria
PtchMDH1	POPTR_004G112800	4	10212570:10215920	1239	412	43.14	8.93	Chloroplast
PtchMDH2	POPTR_008G031700	8	1731727:1736488	1371	456	49.76	7.06	Chloroplast
PtchMDH3	POPTR_017G101900	17	11879289:11882328	1239	412	43.38	8.49	Chloroplast
PtchMDH4	POPTR_017G102000	17	11884524:11887819	1239	412	43.32	8.15	Chloroplast

1 Table 2. Characteristics of the malate dehydrogenase (*MDH*) gene family in *Populus trichocarpa.*

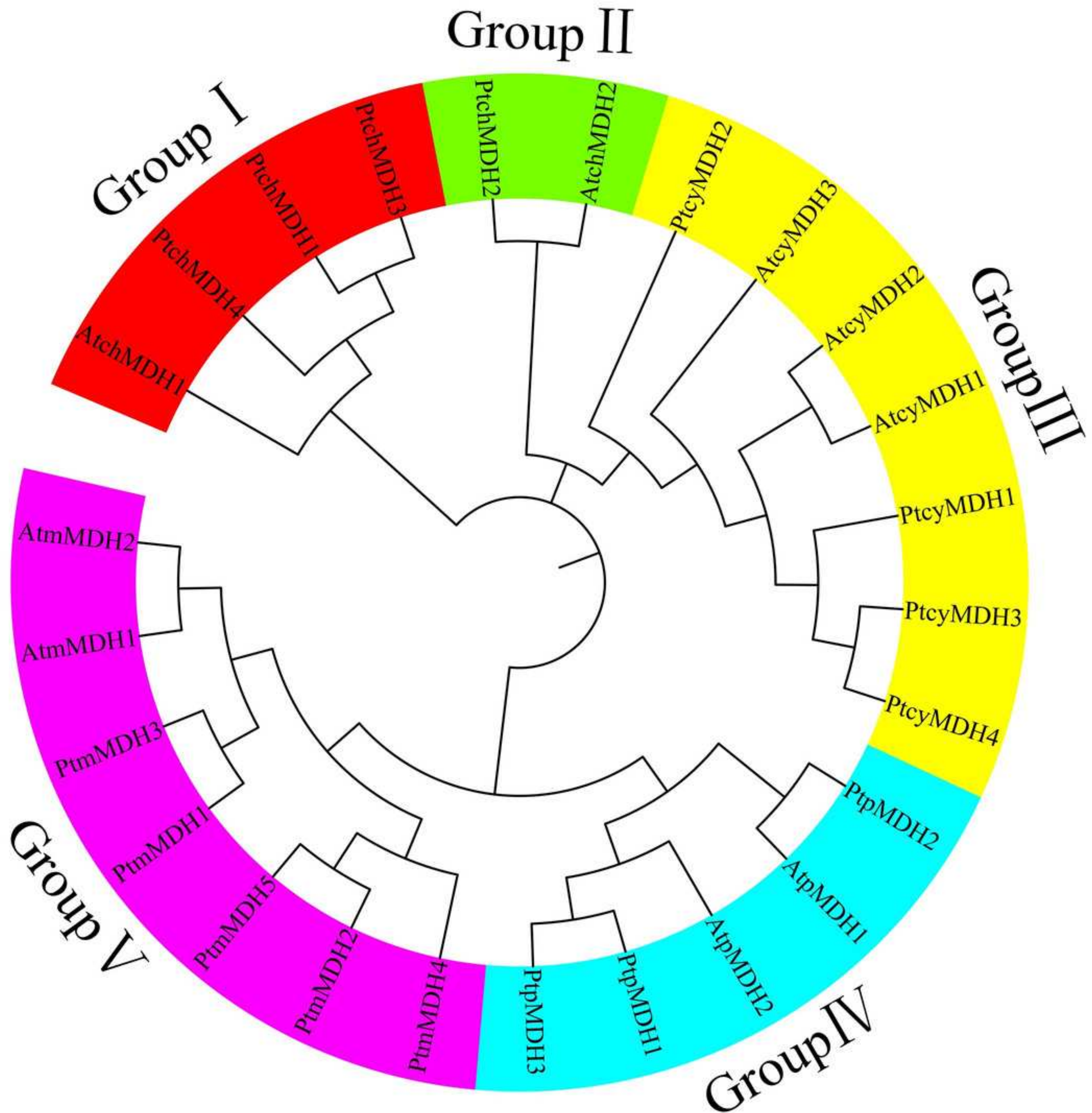
2 Chr: chromosome number; Un: unknown; CDS: coding sequence length (nucleotides); aa: amino acid length;

3 MW: molecular weight; pI: isoelectric point.

Figure 1

Phylogenetic analysis of MDH members in *P. trichocarpa* and *A. thaliana.*

Phylogenetic tree was constructed by MEGA7 with Neighbour-Joining method and bootstrap of 1000 replications. The different colored arcs indicate different groups of MDH proteins.


Group I
Group II
GroupIII
Group IV
Group V
PtchMDH3
PtchMDH1
PtchMDH4
AtchMDH1
PtchMDH2
AtchMDH2
PtcyMDH2
AtcyMDH3
AtcyMDH2
AtcyMDH1
PtcyMDH1
PtcyMDH3
PtcyMDH4
PtpMDH2
AtpMDH1
AtpMDH2
PtpMDH1
PtpMDH3
PtmMDH4
PtmMDH2
PtmMDH5
PtmMDH1
PtmMDH3
AtmMDH1
AtmMDH2

PeerJ Preprints NOT PEER-REVIEWED

Figure 2

Gene structure and conserved motif analysis of *MDH* genes in *P. trichocarpa.*

(A) Phylogenetic tree of 16 PtMDH proteins. (B) Distributions of conserved motifs in *PtMDH* genes. (C) Exon/intron organization of *PtMDH* genes.

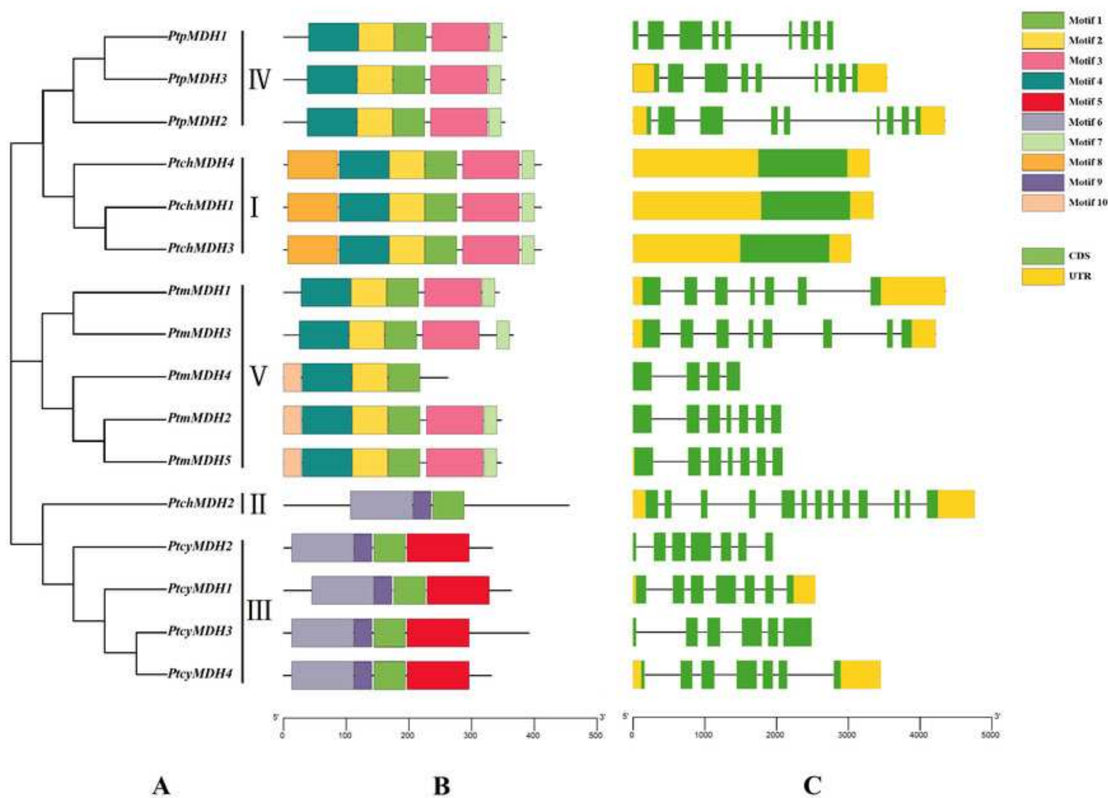
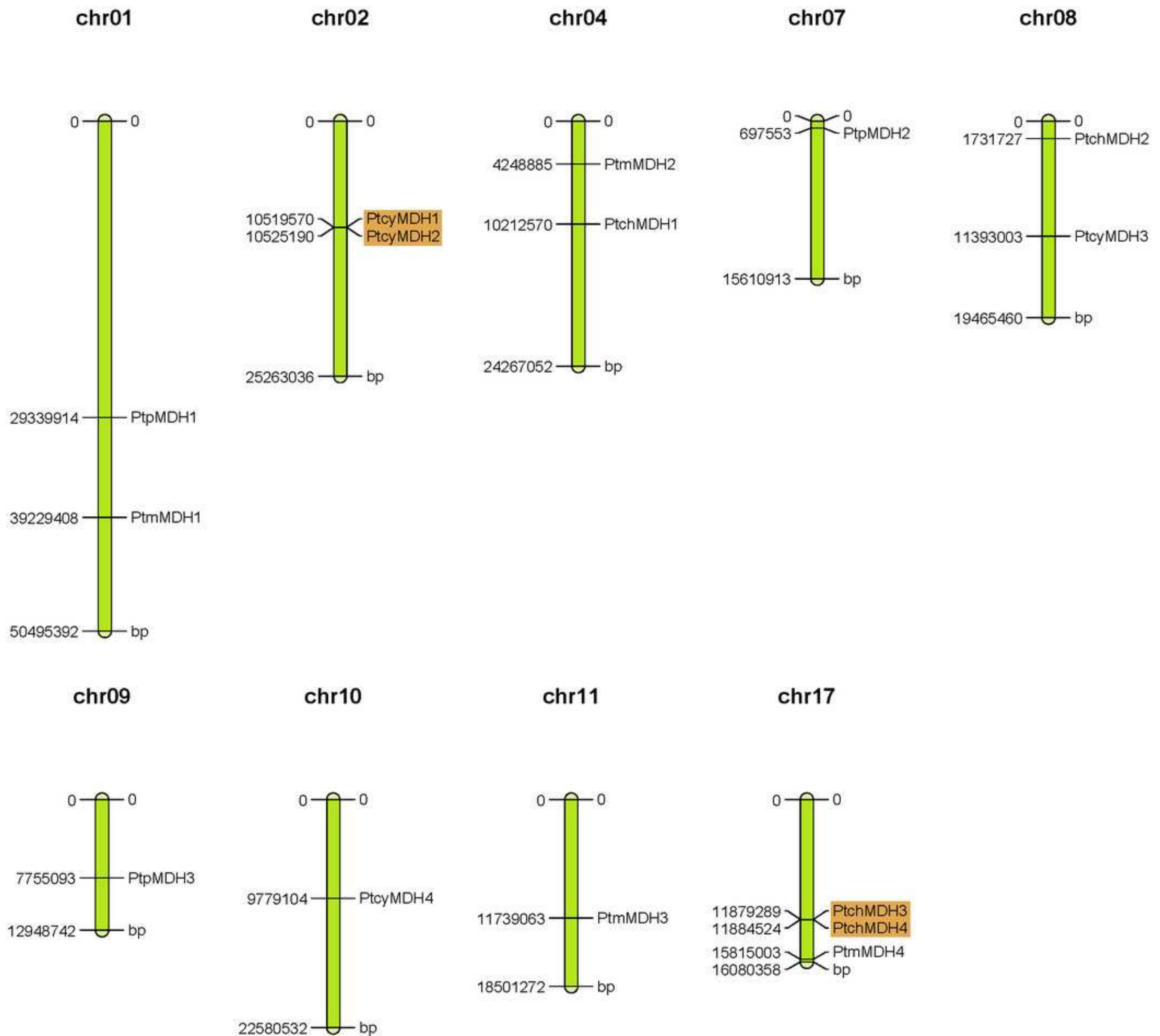


Figure 3

Chromosomal mapping of *MDH* genes in *P. trichocarpa*.

The chromosome numbers and size (bp) are indicated at the top and bottom of each chromosome, respectively. Tandemly duplicated gene pairs are highlighted in yellow.



chr01
0
0
29339914
PtpMDH1
39229408
PtmMDH1
50495392
bp
chr02
0
0
10519570
10525190
PtcyMDH1
PtcyMDH2
25263036
bp
chr04
0
0
4248885
PtmMDH2
10212570
PtchMDH1
24267052
bp
chr07
0
0
697553
PtpMDH2
15610913
bp
chr08
0
0
1731727
PtchMDH2
11393003
PtcyMDH3
19465460
bp
chr09
0
0
7755093
PtpMDH3
12948742
bp
chr10
0
0
9779104
PtcyMDH4
22580532
bp
chr11
0
0
11739063
PtmMDH3
18501272
bp
chr17
0
0
11879289
11884524
PtchMDH3
PtchMDH4
15815003
16080358
PtmMDH4
bp

PeerJ Preprints NOT PEER-REVIEWED

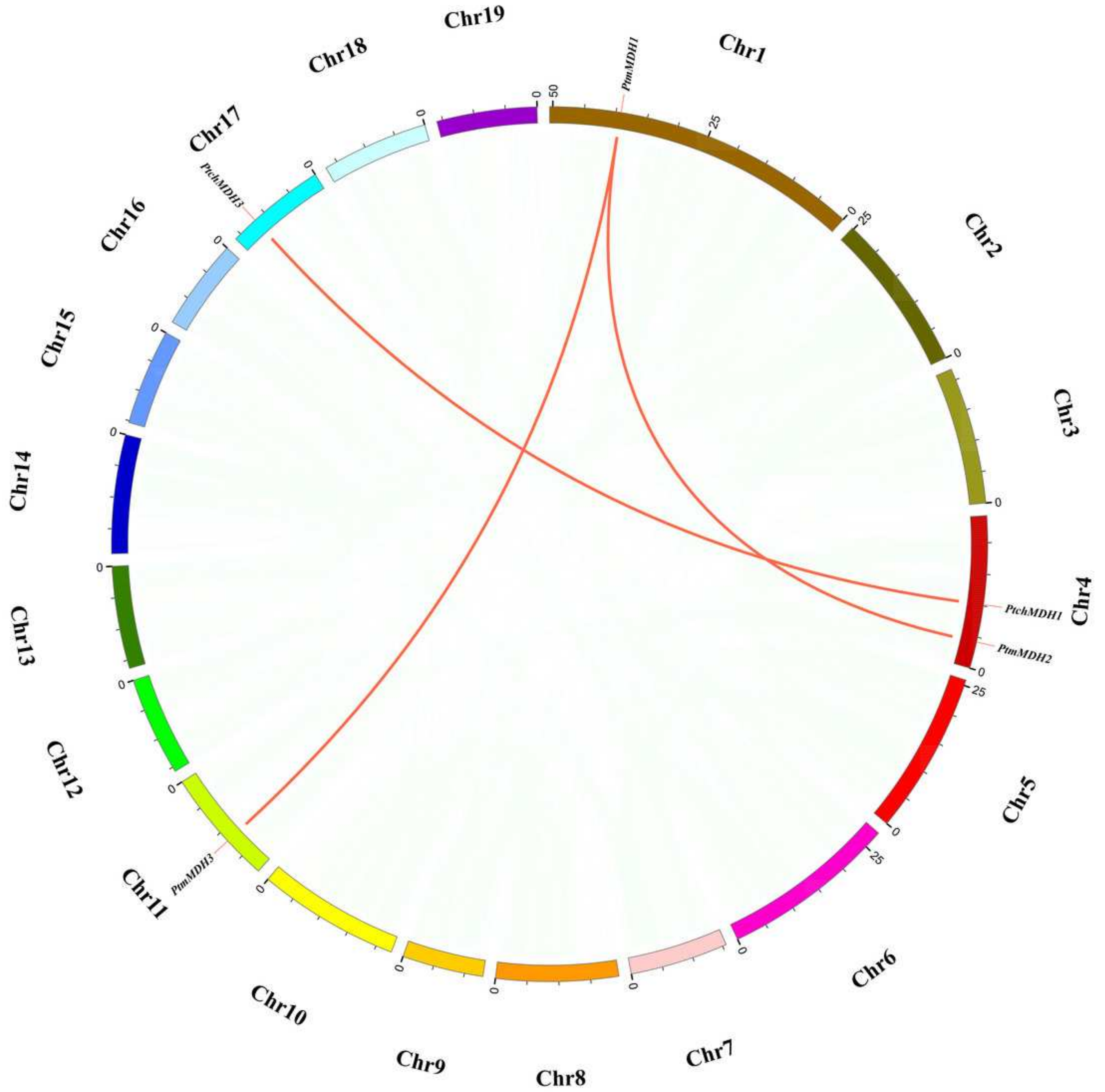
Figure 4

Synteny analysis of the *MDH* gene family in *P. trichocarpa*

Chromosomes were drawn in different colours. Red curves linking *PtMDH* genes represent the segmental duplication events.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27742v1 | CC BY 4.0 Open Access | rec: 18 May 2019, publ: 18 May 2019

PeerJ Preprints NOT PEER-REVIEWED



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27742v1 | CC BY 4.0 Open Access | rec: 18 May 2019, publ: 18 May 2019

PeerJ Preprints NOT PEER-REVIEWED

Figure 5

Expression analysis of *PtMDH* genes in leaf tissues under salt stress.

FPKM values of *PtMDH* genes were transformed by log2. The colour scale represents relative expression levels.

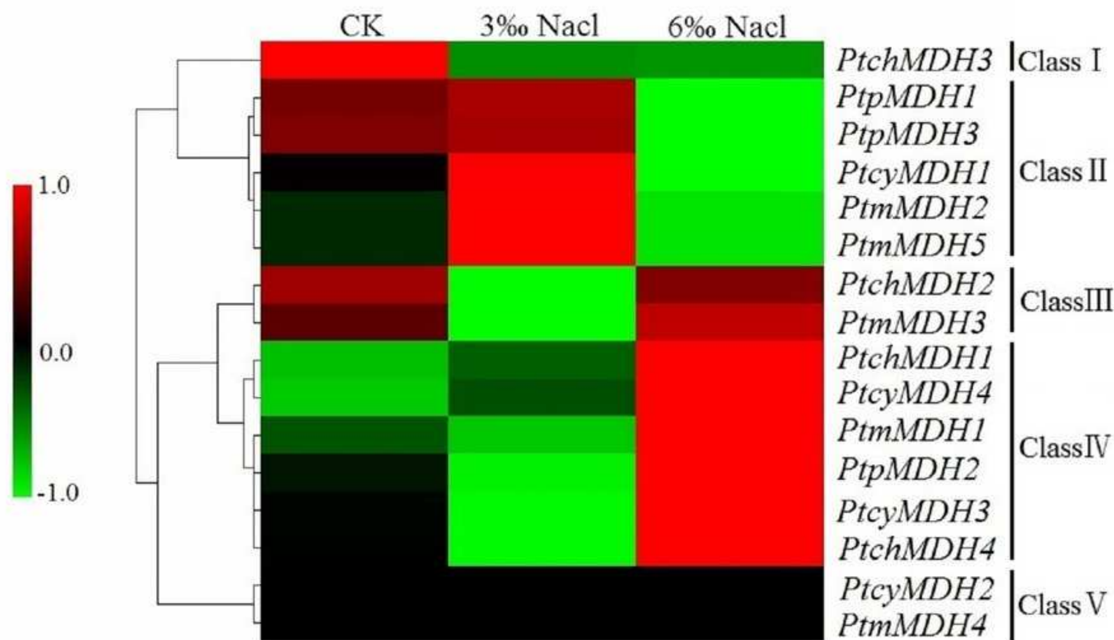


Figure 6

Expression analysis of six *PtMDH* genes in roots under salt stress by qRT-PCR.

Data were normalized to *PtRG5* gene and vertical bars indicate standard error.

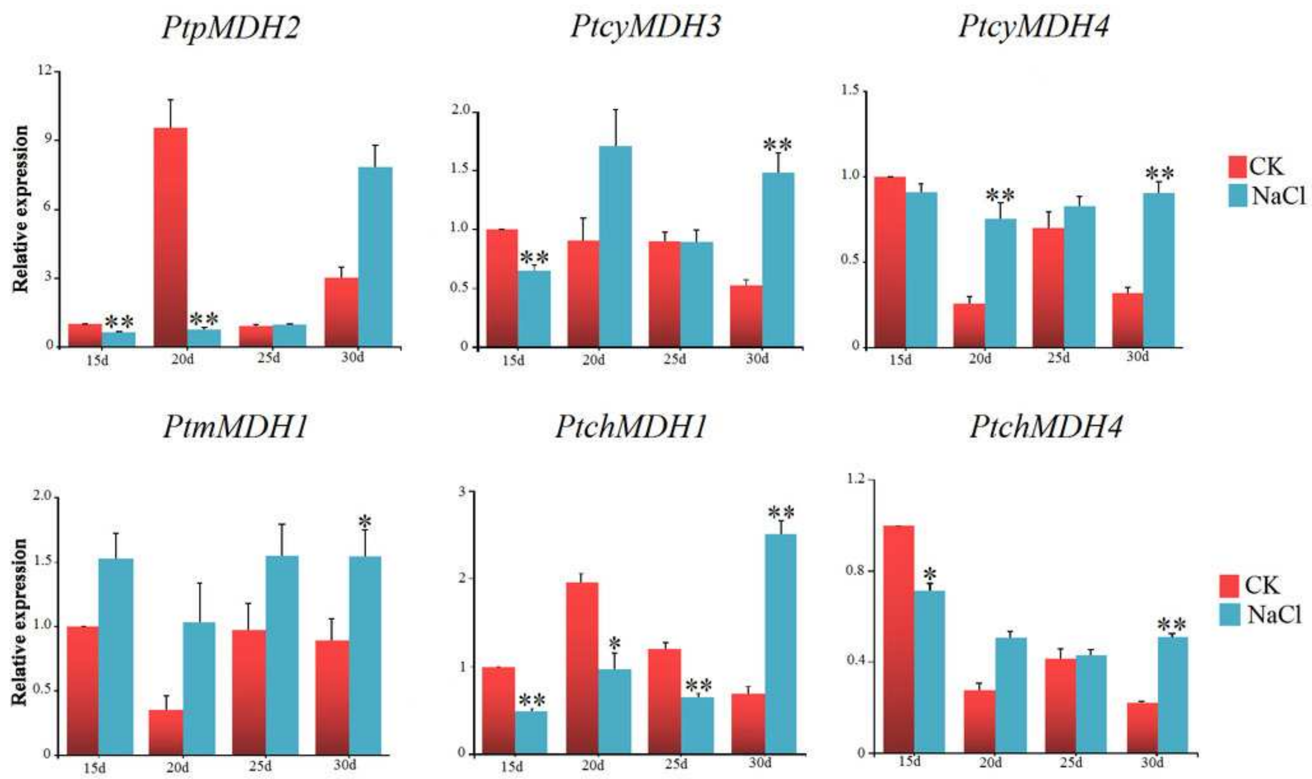
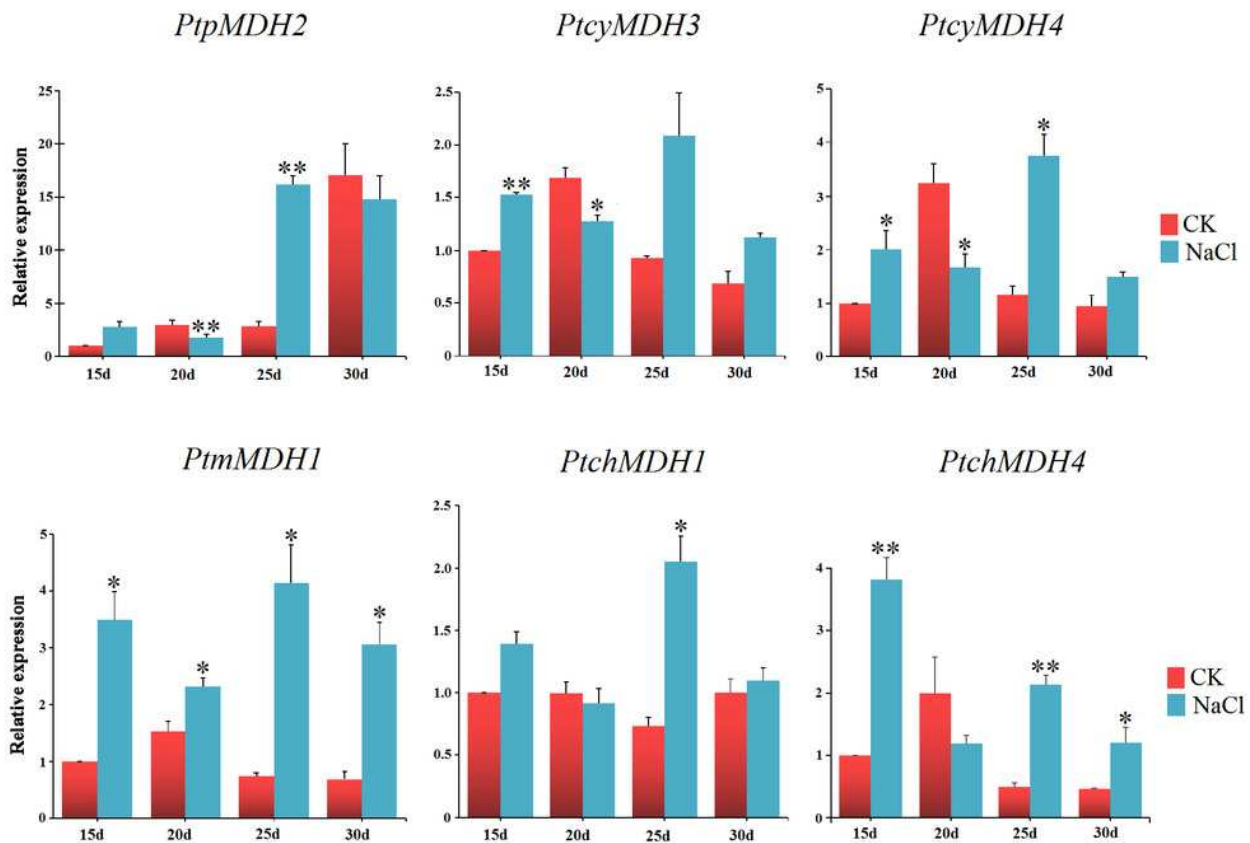


Figure 7

Expression analysis of six *PtMDH* genes in stems under salt stress by qRT-PCR.

Data were normalized to *PtRG5* gene and vertical bars indicate standard error.



PeerJ Preprints NOT PEER-REVIEWED

Figure 8

Expression analysis of six *PtMDH* genes in leaves under salt stress by qRT-PCR.

Data were normalized to *PtRG5* gene and vertical bars indicate standard error.